JMIR Aging

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

Contents

Reviews

Original Papers

Viewpoints

Corrigenda and Addendas

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

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

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

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

[*]all authors contributed equally

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

Abstract

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

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

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

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

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

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

KEYWORDS

gamification; exergaming; enjoyment; satisfaction; older adults

Introduction

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

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

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

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

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

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

Methods

Study Registration

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

Search Strategy

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

Inclusion and Exclusion Criteria

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

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

Study Selection

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

Data Extraction

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

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

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

Risk of Bias

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

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

Main Outcomes

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

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

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

Statistical Analysis

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

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

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

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

Deviations From the Protocol

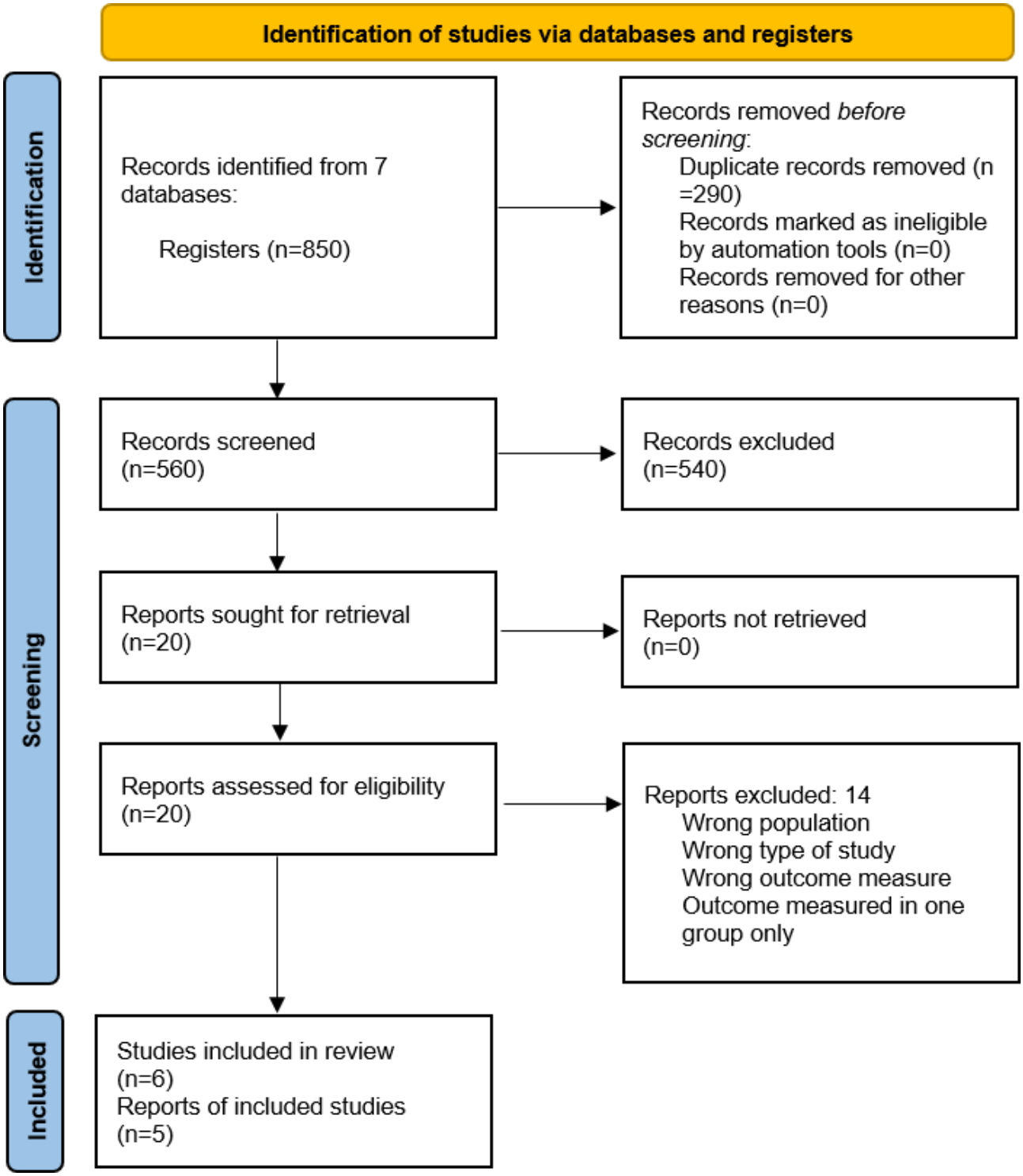
Some analyses foreseen in the protocol registered in PROSPERO could not be performed in this review. Secondary outcomes were foreseen to meta-analyze the adherence to exercise through video games and the adverse effects that these interventions could have; however, this analysis could not be performed,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

Results

Study Selection

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

Figure 1. Flow diagram in this review.



Characteristics of the Included Studies

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

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

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

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

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

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

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

Table . General study characteristics.

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

[a]RCT: randomized controlled trial.

[b]M: men.

[c]W: women.

[d]mCI: mild cognitive impairment.

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

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

[a]AR: augmented reality.

[b]USQ: User Satisfaction Questionnaire.

[c]N/A: not applicable.

[d]PACES: Physical Activity Enjoyment Scale.

[e]VR: virtual reality.

[f]IMI: Intrinsic Motivation Inventory.

[g]RPE: Rate of Perceived Exertion.

Risk of Bias in the Included Studies

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

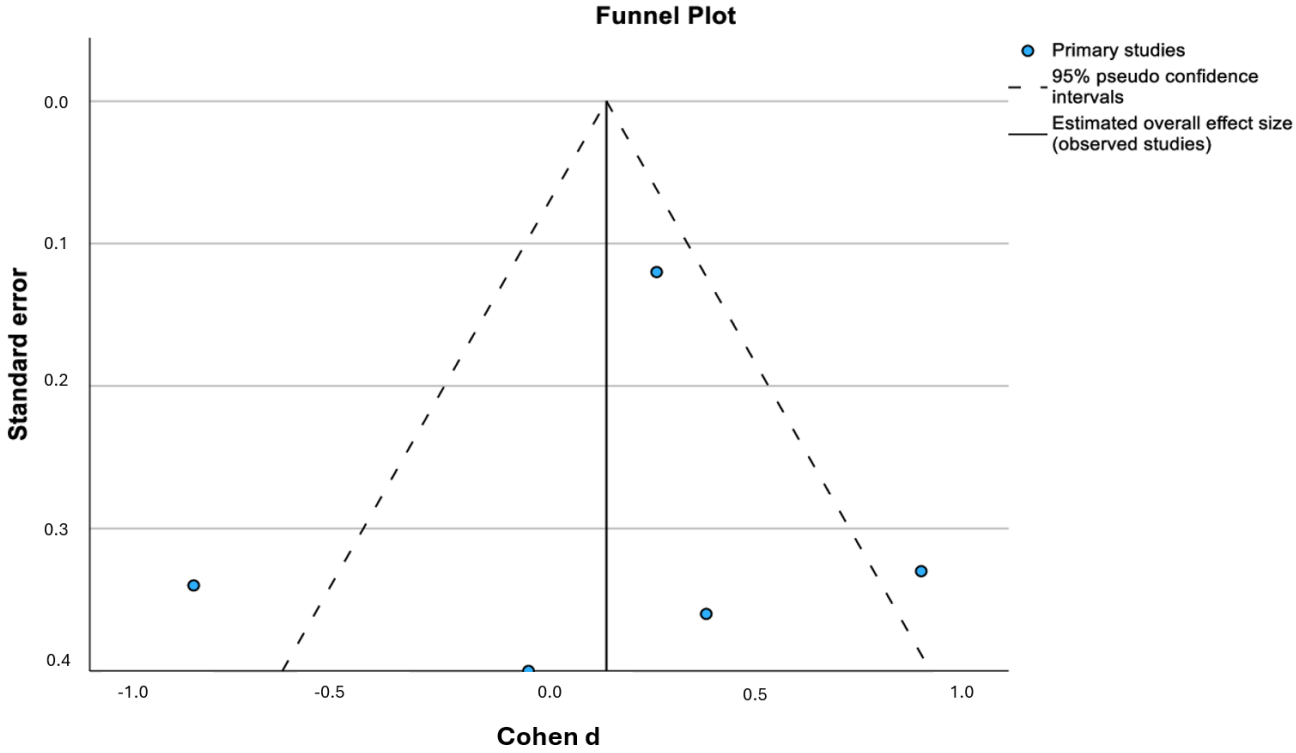
The Egger regression–based test was conducted to evaluate the presence of publication bias. The results indicated that the intercept was not significantly different from 0 (intercept=2.985; P=.83), suggesting no evidence of small-study effects.

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

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

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

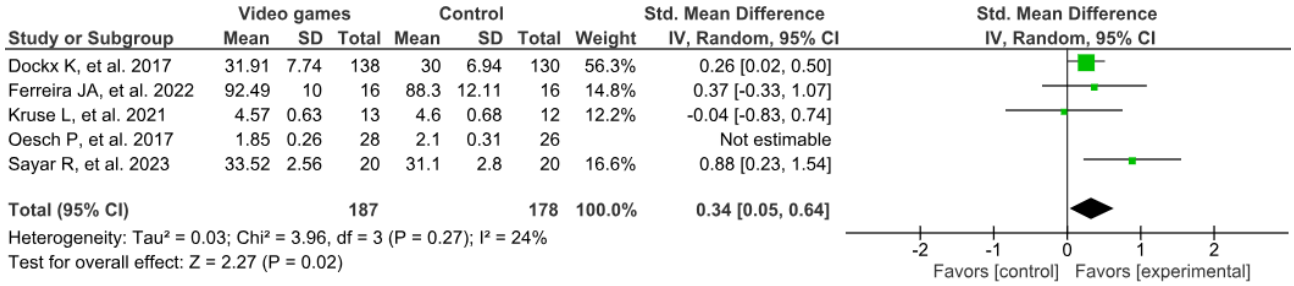
Figure 3. Funnel plot of the included studies.



Quantitative Analysis

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

Figure 4. Forest plot of the overall enjoyment.



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

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

Table . Subgroup analysis.

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

[a]SMD: standardized mean difference.

[b]N/A: not applicable.

Discussion

Principal Findings

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

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

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

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

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

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

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

Limitations

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

Clinical Implications

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

Conclusion

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

Acknowledgments

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

Conflicts of Interest

None declared.

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

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

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

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

References

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

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

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

Abbreviations

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

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

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

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

Methods

Study Design and Participants

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

Ethical Considerations

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

Participants

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

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

Randomization and Masking

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

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

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

Intervention Program: GrandMove

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

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

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

Measures

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

Primary Outcomes

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

Secondary Outcomes

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

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

Statistical Analysis

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

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

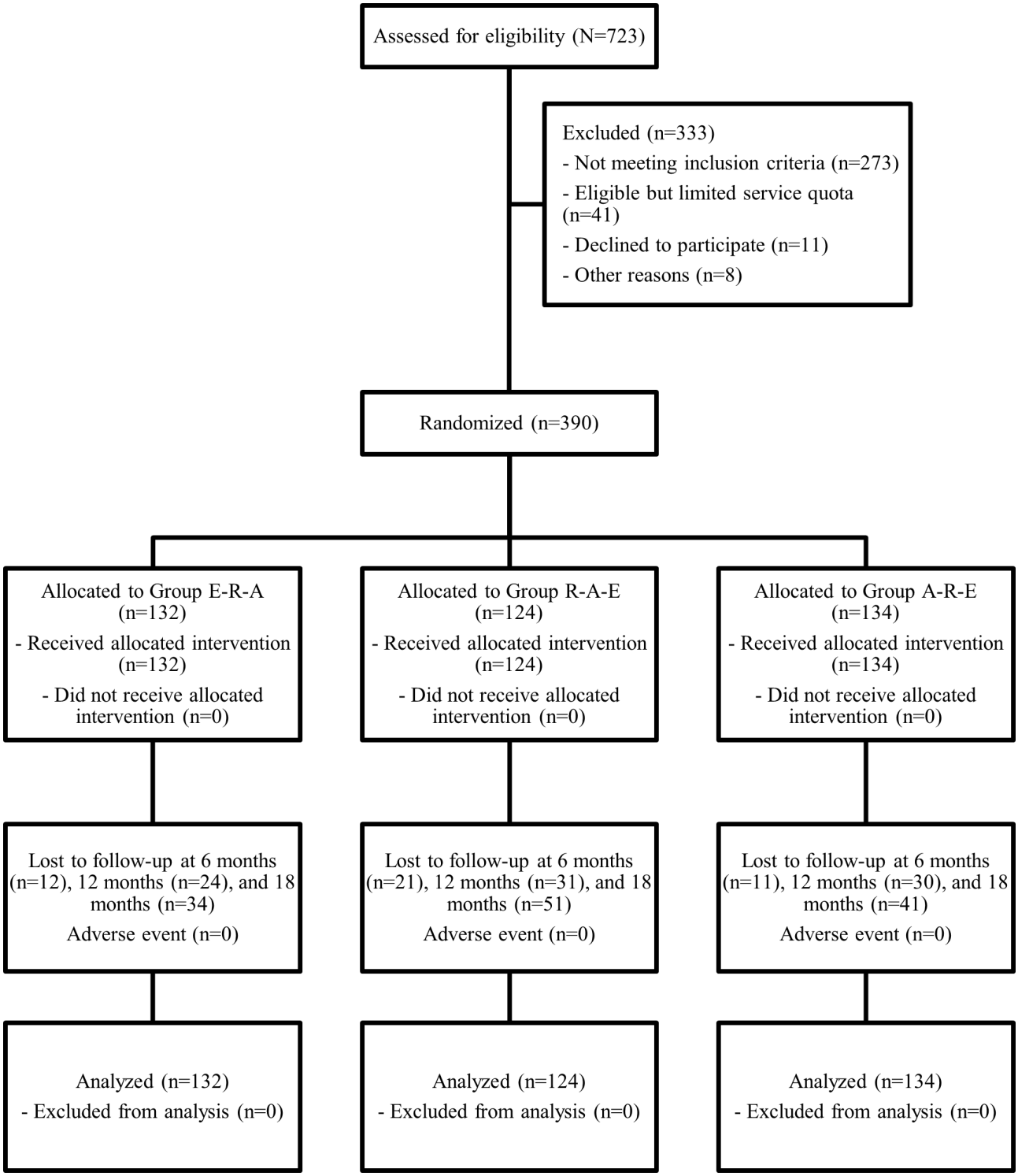
further adjust for other baseline covariates that were found to be significantly different between groups. To examine the effectiveness of 6-month aerobic or resistance training, outcomes in the A-R-E group and R-A-E group were compared to the E-R-A group. To investigate the order effect in aerobic and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

Results

Sample Characteristics

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

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



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

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

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

[a]E: lifestyle education.

[b]R: resistance training.

[c]A: aerobic training.

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

[e]SPPB: Short Physical Performance Battery.

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

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

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

[i]LSNS: Lubben Social Network Scale.

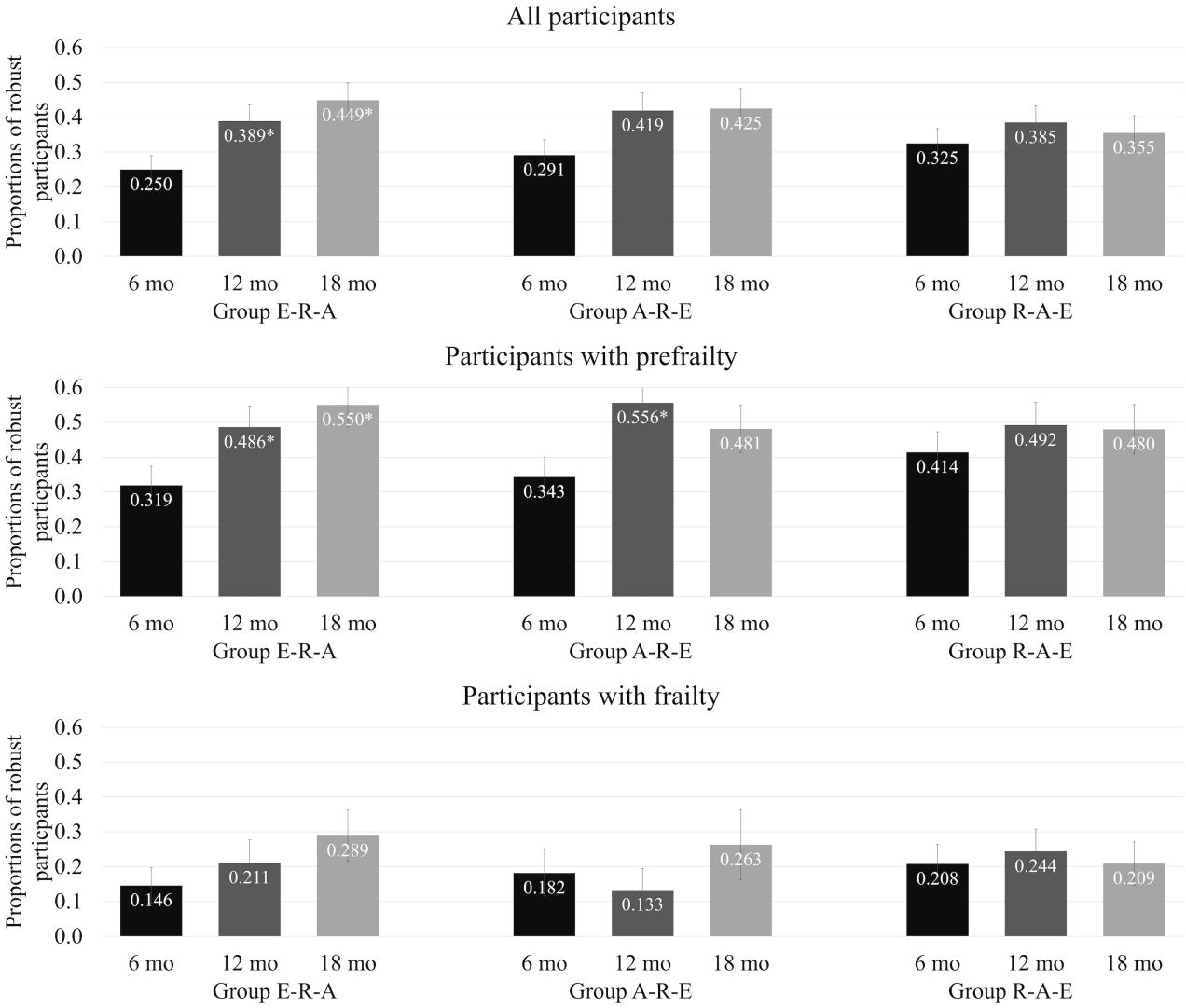
[j]PSQI: Pittsburgh Sleep Quality Index.

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

Change in Frailty Status

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

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



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

Aerobic Versus Resistance Training: 6-Month Outcomes

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

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

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

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

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

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

[c]SPPB: Short Physical Performance Battery.

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

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

[f]LSNS: Lubben Social Network Scale.

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

[h]PSQI: Pittsburgh Sleep Quality Index.

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

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

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

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

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

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

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

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

Maintenance of Exercise Effect and Habit Formation

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

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

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

Discussion

Principal Findings

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

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

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

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

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

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

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

Limitations

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

Conclusions

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

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

Acknowledgments

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

Authors' Contributions

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

Conflicts of Interest

None declared.

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

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

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

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

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

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

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

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

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

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

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

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

References

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

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

Abbreviations

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

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

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

Methods

Study Population

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

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

Ethical Considerations

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

Assessment of Lifestyle Factors

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

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

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

Ascertainment of Outcomes

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

Assessment of Covariates

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

Statistical Analysis

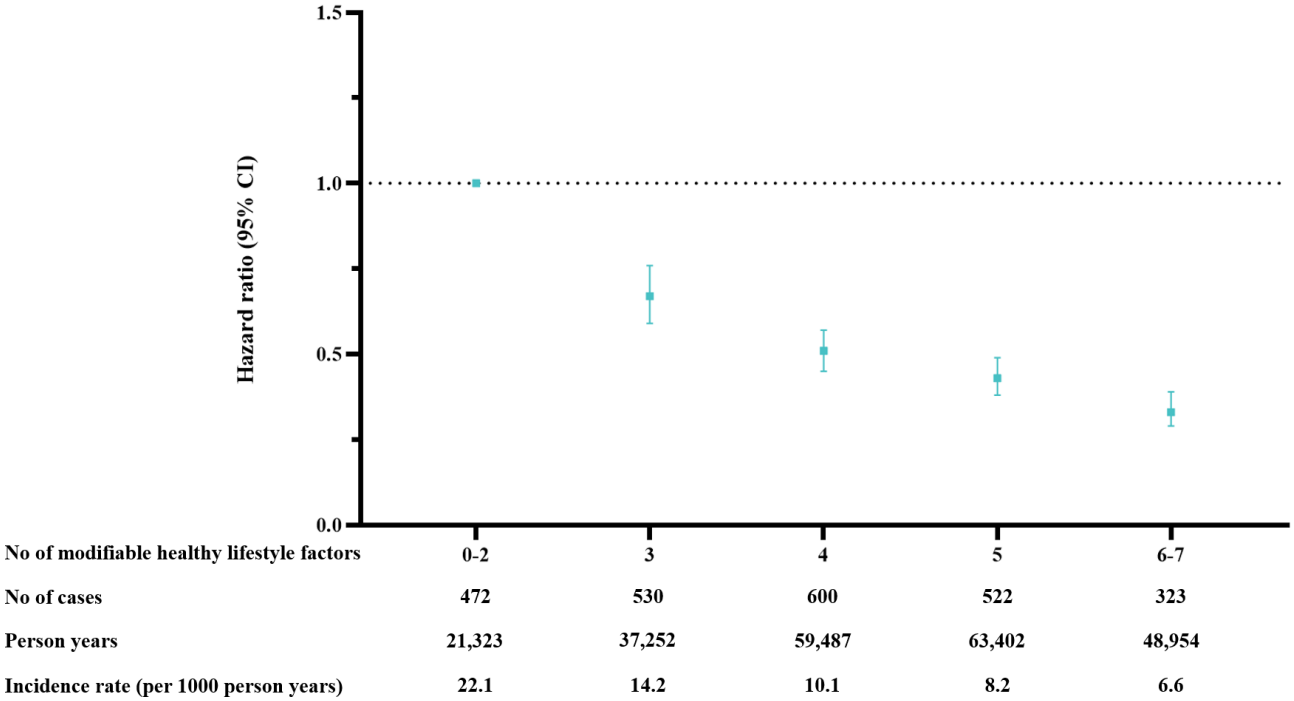
Baseline characteristics are presented as the mean (SD) if normally distributed or median (IQR) if nonnormally distributed for continuous variables and frequency (%) for categorical variables. The distribution of quantitative variables was evaluated using both graphical assessments (ie, histograms and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

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

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

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

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



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

Results

Population Characteristics

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

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

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

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

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

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

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

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

[a]HR: hazard ratio.

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

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

Association of Lifestyle Score With Risk of All-Cause Mortality

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

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

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

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

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

[a]HR: hazard ratio.

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

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

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

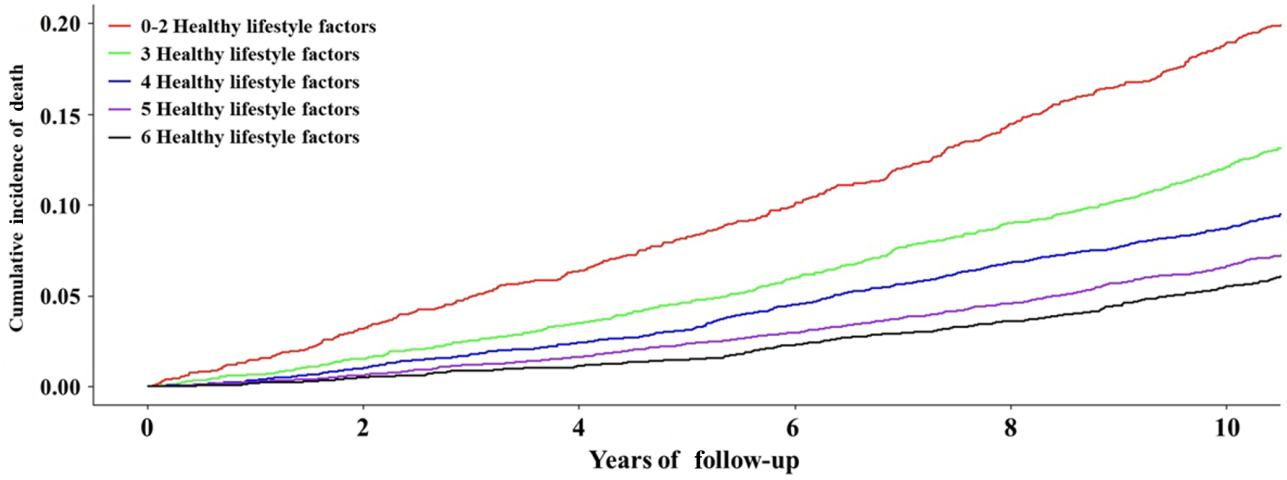


Table . Relative and population-attributable risks of all-cause mortality for groups defined by combinations of modifiable healthy lifestyle factors.

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

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

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

[c]PAR%: population attributable risk percent.

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

Association of Lifestyle Score With Risk of Cause-Specific Mortality

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

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

Discussion

Statement of Principal Findings

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

Comparison With Previous Studies

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

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

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

Potential Mechanisms

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

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

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

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

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

Clinical and Research Implications

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

Strengths and Limitations

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

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

Conclusions

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

Acknowledgments

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

Authors' Contributions

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

Conflicts of Interest

None declared.

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

References

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

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

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

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

Abbreviations

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

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

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

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

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

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

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

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

KEYWORDS

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

Introduction

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

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

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

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

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

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

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

Methods

Participants

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

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

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

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

Ethical Considerations

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

Study Design

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

Intervention

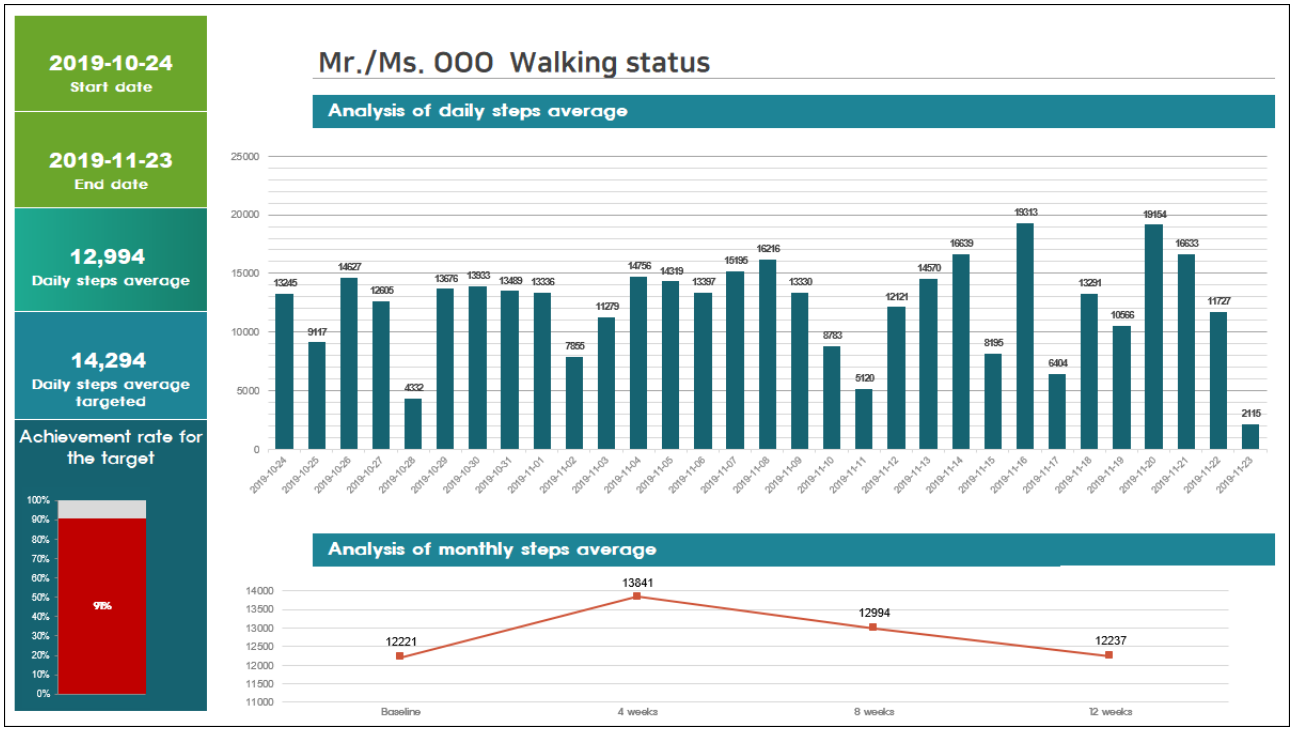
Our 12-week protocol of intervention was designed for use in a hospital setting. Physicians, nurses, psychologists, physical or occupational therapists, and any other trained paramedic professionals can use this program. In this study, a clinical psychologist and occupational therapist alternatively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

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

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

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

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



Outcome Assessment and Measures

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

Primary Outcome

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

Secondary Outcomes

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

Statistical Analyses

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

Results

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

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

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

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

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

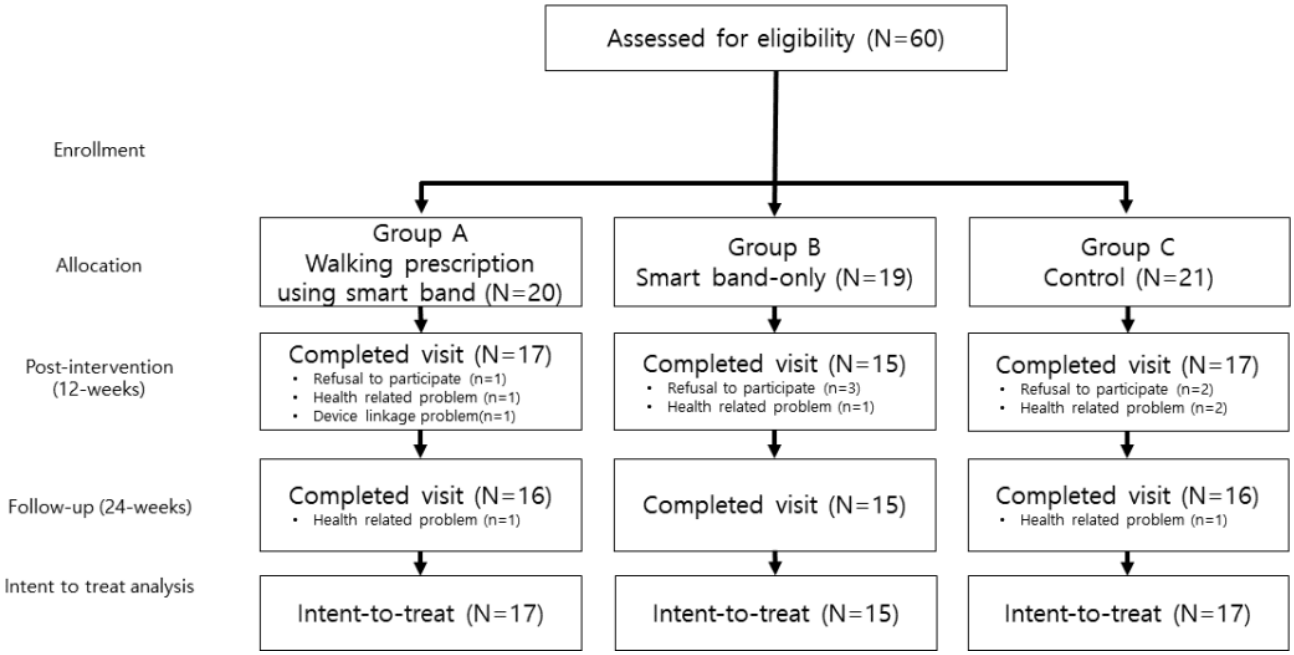


Table . Baseline characteristics of participants.

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

[a]SPPB, Short Physical Performance Battery.

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

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

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

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

[a]Control as a reference group.

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

[c]SPPB, Short Physical Performance Battery.

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

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

[f]Baseline as a reference time.

[g]$P<.001$.

[h]$P<.05$.

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

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

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

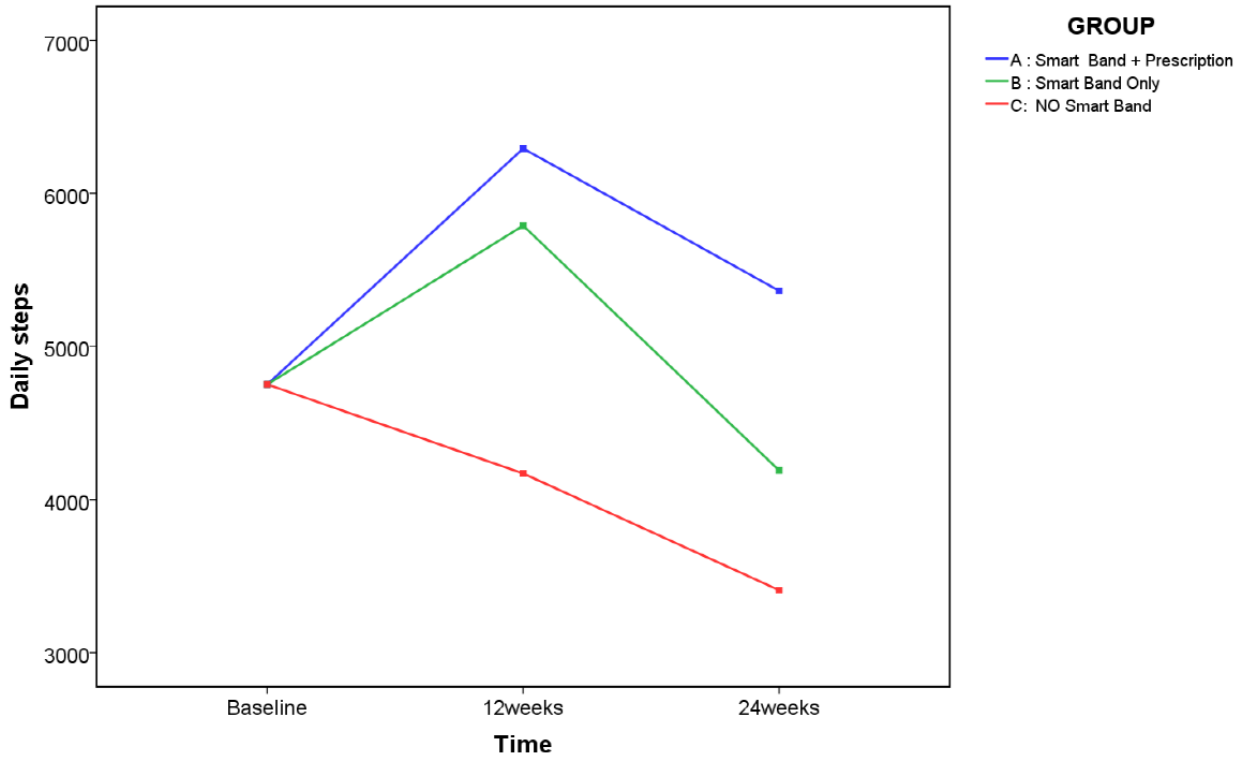


Table . Intra-group changes in daily steps and cognitive function from baseline to 12 weeks and 24 weeks follow-up.

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

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

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

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

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

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

Discussion

Principal Findings and Comparison With Previous Works

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

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

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

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

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

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

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

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

Strengths

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

Limitations

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

Conclusions

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

Acknowledgments

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

Authors' Contributions

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

Conflicts of Interest

None declared.

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

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

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

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

References

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

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

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

Abbreviations

ANOVA: analysis of variance

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

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

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

Impact of a Light Volleyball Intervention Program on Improving Physical Attributes of Older Adults in Hong Kong: Preliminary Study of a Randomized Controlled Trial

Ka Man Leung[1], BSci, MSocSci, MSci, PhD; Yuchen Shi[2], PhD

[1]Department of Health and Physical Education, Faculty of Liberal Arts and Social Science, Education University of Hong Kong, Tai Po, China (Hong Kong)

[2]School of Liberal Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China

Corresponding Author:
Yuchen Shi, PhD
School of Liberal Arts, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China

Abstract

Background: Physical inactivity, which increases the risk of chronic diseases in older adults, is prevalent among older adults in Hong Kong. To address this problem, the Hong Kong government has been proactively promoting active aging.

Objective: Following the World Health Organization's strategy to prevent chronic diseases in older adults and aligning with the global goal of active aging, this study evaluated the effects of a 16-week light volleyball (LVB) intervention program on the physical health of older adults in Hong Kong.

Methods: A total of 276 participants aged ≥60 years were recruited and randomly assigned to 1 of 3 groups: an LVB intervention group, a Taichi control group (ie, with light physical activity), and a control group. Tests on components of fitness were conducted before and after the intervention.

Results: Participants from the LVB intervention group exhibited significant improvements in lower body strength ($F_{2,272}$= 7.23, P=.001, η^2=.05), agility ($F_{2,272}$= 6.05, P=.003, η^2=.043), and dynamic balance ($F_{2,272}$= 9.41, P=.001, η^2=.065) when compared with those from the Taichi active control group and control group.

Conclusions: To promote active aging among older adults in Hong Kong, the findings of this preliminary study, along with forthcoming follow-up tests, will provide health specialists and practitioners with valuable insights regarding the health benefits of the LVB community program for older adults.

Trial Registration: Chinese Clinical Trial Register ChiCTR1900026657; https://www.chictr.org.cn/showprojEN.html?proj=44350

(JMIR Aging 2025;8:e62886) doi:10.2196/62886

KEYWORDS

adapted physical activity; older adults; gerontology; geriatrics; older; aging; randomized controlled trial; volleyball; light volleyball; intervention; sports; physical activity; exercise; physical attributes; RCT; controlled trial

Introduction

Given the low birth rate and increasing life expectancy in Hong Kong, the trend toward an aging population is expected to persist in the coming decades in this region. The number of individuals aged ≥65 years is projected to rise substantially, reaching 1.89 million by 2019 and 30.5 million by 2069 [1]. This continual demographic shift highlights several societal challenges, including a decline in the working population and an increased burden on Hong Kong's social welfare and health care systems. With an expanding share of older adults in the population, public health expenditure in Hong Kong is forecasted to increase by 394%, from US $4.9 billion in 2004 to US $24.4 billion in 2033, accounting for 5.5% of Hong Kong's total GDP in 2033 [2]. The simultaneous decrease in the working population and increase in social welfare expenditures for the aging population are expected to create severe financial strain for the Hong Kong government in the coming decades.

Although the benefits of physical activity have been well-documented [3], approximately 40% of older adults aged 60 years or older in Hong Kong engage in insufficient physical activity [4]. More than 50% of older adults in Hong Kong have overweight or obesity [5], and 75% have one or more chronic illnesses [6]. Additionally, older adults in Hong Kong spend an average of 7.5 hours sitting and 6.7 hours lying down daily [7]. Spiteri et al [8] identified various barriers to physical activity for older adults, including perceptions of negative consequences (ie, pain, risk of injury, and fear of falling), and a lack of knowledge and skills. Similar barriers to physical activity have been reported among older adults in Hong Kong. To address

these challenges, promoting active aging through physical activity is essential [9]. The World Health Organization (WHO) [10] introduced the "Active Ageing: A Policy Framework," warning that failure to address the aging population crisis could lead to the collapse of health care and social welfare systems. The Hong Kong government interprets active aging as achieving full physical, social, and mental well-being [6]. Since 2001, it has developed a healthy aging report outlining strategies to promote active aging, including physical activity initiatives. For instance, the Hong Kong government installed 150 new fitness equipment sets for older adults across Hong Kong between 2017 and 2018 to encourage active aging [11].

Meta-analyses have demonstrated that physical activity interventions have a significant effect on improving balance, reducing falls [12], and enhancing physical function [13] in older adults. Both supervised resistance and aerobic physical activity interventions have demonstrated positive effects on physical function outcomes. Future research directions include (1) increasing the number of physical activity interventions conducted in Asia, (2) aligning intervention programs with the WHO's recommendation of 150 minutes of physical activity per week, (3) using more rigorous randomized controlled trial designs, and (4) implementing supervised physical activity interventions with larger sample sizes in Asia [14,15]. Recent reviews have emphasized the need to examine the effects of sports interventions on older adults' health, particularly among Chinese older adults. Wong et al [16] reviewed 371 studies from the past 15 years on the effects of physical activity on older adults' physical health, and they recommended exploring the health benefits of newly emerging sports, such as light volleyball, in future research [16,17]. Franco et al [17] qualitatively reviewed the literature on older adults' perceptions of physical activity participation, suggesting that older adults prefer physical activity programs that are (1) professionally instructed, (2) group-based and conducive for peer interaction, and (3) highly accessible, with affordability being a key factor influencing the willingness to participate [17]. Similar findings were reported by Van Dyck et al [18], who highlighted older adults' preferences for innovative physical activity activities, such as aqua fitness or light volleyball, over traditional options, such as walking and cycling, in intervention programs.

Light volleyball (LVB) is a newly adapted physical activity derived from traditional volleyball. Compared with traditional volleyball, LVB uses balls with a larger circumference (80 - 83 cm vs 65 - 67 cm) and lighter weight (150g vs 250g), allowing the balls to travel at a slower speed. These features make LVB more accessible, particularly for individuals with reduced physical capacity due to age-related decline (eg, older adults). Additionally, the LVB court is smaller (similar in size to a badminton court), and the net is set lower (1.8 m), further reducing the physical demands on participants. Studies from China have reported physical health benefits associated with regular LVB practice; however, most of these studies did not have standardized fitness measurements and control groups. In 2020, the first author and her team conducted a quasi-experimental intervention to examine the effects of LVB on the physical and psychological health attributes of 78 older adults aged ≥60 years. The results indicated significant improvements in physical (ie, lower and upper body muscle strength, agility, balance, and aerobic endurance) and psychological (ie, physical activity enjoyment) attributes among participants in the LVB group when compared with the control group [19]. Furthermore, improvements in upper body muscle strength, aerobic endurance, and physical activity enjoyment were significantly more pronounced in the LVB group than in the active control group, which participated in Rouliqiu. That study highlighted the effectiveness of LVB intervention programs in enhancing both physical and psychological health attributes in older adults. Despite these promising findings, that study had notable limitations. First, its sample size was small, with only 62 participants completing the screening tests, pre- and postintervention functional tests, and data analysis. Second, participants were not randomly assigned to groups, which could have introduced selection bias.

Building on the positive results of the LVB pilot study and the prioritization of resource allocation for promoting active aging in Hong Kong, the first author and her team secured Research Impact funding amounting to US $0.95 million for further research in this area. This funding supported an investigation into the effectiveness of an LVB intervention on the physical and psychological health attributes of older adults in Hong Kong using both quantitative and qualitative methods and examining a larger sample size of approximately 300 participants. In this study, we present the preliminary results of this LVB intervention in our quantitative arm using the results from the pretest and posttest. In the current study, Taichi was selected as the intervention for the active control group because both Taichi and LVB are whole-body exercises originating from China and suitable for older adults [20,21]. Compared with the team-based LVB, Taichi is an individual exercise, which may lead to differing effects on older adults' health and quality of life [22]. Prior research suggests that older adults with greater social support are more likely to continue exercising regularly [23,24]. Community-based group physical activity interventions with increased social support have also been associated with greater beneficial effects and program adherence. Additionally, Taichi is a popular activity among Chinese older adults [25]. With LVB gaining traction in Hong Kong, Taichi serves as a relevant comparison, providing valuable insights into physical activity promotion among Hong Kong's older adult population. This study aimed to assess the effects of a 16-week LVB intervention compared with Taichi and a control group on two key outcomes in Chinese older adults aged ≥65 years: (1) functional fitness and (2) balance. The hypothesis was that both LVB and Taichi would result in significant and comparable improvements in these physical health attributes relative to the control group.

Methods

Study Design

A randomized controlled trial design was used to assess the effects of the LVB intervention on participants' physical health outcomes. Following the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1) [26], detailed information about the intervention can be found in our

previously published study protocol (trial registration number ChiCTR1900026657) [19]. The LVB group was compared against an active control group (ie, Taichi) and a control group. Participants were randomly assigned into LVB group, Taichi group, and control group in 1:1:1 ratio.

Study Intervention

The intervention program was conducted from mid-2021 to early-2023, with data collection performed at pretest (immediately before the intervention) and posttest (immediately after the intervention). The preliminary results presented here are based on pretest and posttest data only. Participants in the LVB and Taichi groups attended a 16-week training program consisting of two 90-minute sessions per week following the pretest. This intervention duration aligns with the guidelines of the Centers for Medicare and Medicaid Services of the US Department of Health and Human Services, which suggest that the health benefits of physical activity for older adults can be observed within 1 to 3 months after the start of a program [27]. In contrast, the control group was instructed to continue their usual daily activities while participating in monthly social gatherings (eg, health workshops) to control for psychosocial effects.

Participants

Participants were eligible for recruitment if they (1) were aged ≥60 years, (2) lived independently, (3) had no cognitive impairment, (4) had not participated in physical activity programs for 2 consecutive years prior to the program, and (5) had passing scores on the Abbreviated Mental Test (AMT) and Timed-up-and-go (TUG) test [28,29]. Specifically, individuals had to obtain a score of at least 6 out of 10 on the AMT to demonstrate sufficient cognitive capability and complete the TUG test within 20 seconds. Participants were ineligible for recruitment if they had steady hypertension (≥160/90 mmHg), arthritis, or neurological disorders.

Recruitment and Procedures

Participants were recruited through informational sessions conducted by the research team and advertisements placed in local neighborhood elderly centers. All participants were informed about the confidentiality of their personal data collected for the study and were assured they could withdraw from the program at any time. The research team sought consent from each participant before collecting data and proceeded immediately to conduct the AMT and TUG test to screen unqualified participants. After participants passed the screening tests, they completed the following steps: (1) questionnaires, (2) sociodemographic questions, (c) measurements of weight and body fat percentage using the Tanita machine (TBF-410GS), and (d) a functional fitness test based on the work of Leung et al [19,20]. After the pretest, an independent researcher used a computer-generated random number system to assign participants to the intervention conditions. Data collection and entry personnel were blinded to the group assignments of participants. The intervention program commenced the following week in sports complexes at community centers or community halls in Hong Kong and lasted for 16 weeks (approximately 3.5 months). The research team arranged the posttest for participants within 7 days after they completed the intervention program. All participants received a supermarket cash voucher worth US$ 12.84 as an incentive for their participation.

Ethical Considerations

The study was reviewed and approved by the Education University of Hong Kong's Research Ethics Committee (approval number E2022-2023-0013).

Measures

Functional Fitness

The research team used the Senior Fitness Test Manual to assess the physical attributes of older adults [30]. The test comprises 6 items: the chair stand test (lower body strength), arm curl test (upper body strength), chair sit-and-reach test (lower body flexibility), back-scratch test (upper body flexibility), 8-foot up-and-go test (agility and balance), and 2-minute step test (aerobic endurance). These have been demonstrated to be reliable, with intraclass correlation coefficients ranging between .80 and .98 in participant trials. Their validity has been supported through content, criterion-related analyses, and construct validation, including comparisons of senior fitness test scores with other established measures, such as treadmill VO2 testing [31]. Higher scores in the chair stand test (repetitions), arm curl test (repetitions), chair sit-and-reach test (cm), back-scratch test (cm), and 2-minute step test (repetitions) indicate higher levels of lower and upper body strength, flexibility, and aerobic endurance; the converse is the case for the 8-foot up-and-go test (seconds). For the flexibility assessments, negative scores reflect an inability to reach the toes during the chair sit-and-reach test or to make hand contact in the back-scratch test.

Balance Test

The Balance System SD (BBS-SD, 950 - 441 model) was used to measure the dynamic balance of older adults in the current study. Participants were asked to stand upright at the center of the platform while observing a screen situated 30 cm (approximately 11.1 in) in front of them. They completed three 20-second trials with 10-second breaks in between. The results of 3 trials were collected, and the mean values were recorded by the research team. The findings of previous studies support the reliability and validity of this balance test [32,33]. The measurement index and overall stability index generated by the Biodex balance system were included in the current study. These indices measured the participants' dynamic balance, specifically assessing their balance fluctuations across multiple axes. Higher values indicated greater deviations and poorer balance control.

Data Analysis

Data were analyzed using SPSS 27.0 software (IBM Corp). Descriptive statistics, specifically mean and SD values for continuous variables and frequencies and percentages for categorical variables, were used to describe the data. Preliminary checks were conducted to ensure that the assumptions of normality and homogeneity of variance were met. One-way ANOVA was performed to evaluate baseline differences between the groups. To examine the effect of the 16-week intervention program on physical attributes, a series of analysis of covariance tests were conducted, comparing the 3 groups

(LVB, Taichi, and control groups) at 2 time points (pre- and posttest). Body mass was used as a covariate because of its correlation with the outcome measures [34]. Partial eta squared (η^2), *P* values, and contrast *t*-statistics were calculated. Planned contrasts were used to analyze the differences between the LVB and control groups, and between the LVB and Taichi groups. Cohen *d* was calculated as a measure of effect size, with thresholds of 0.1 for a small effect, 0.3 for a medium effect, and 0.5 for a large effect [35]. Statistical significance was set at $P<.05$.

Results

Overview

Figure 1 presents the recruitment process for the current study. At the start of the intervention, 334 older adults from 7 elderly centers were enrolled in the program and completed the pretest. With simple randomization, there were 122 older adults in LVB group, 113 in Taichi group, and 99 in control group. A total of 3 participants (2 from the LVB group and 1 from the Taichi group) withdrew from the program because of personal reasons, and 7 others (2 from the LVB group, 1 from the Taichi group, and 3 from the control group) withdrew without providing any reason. Among the remaining 318 participants, 42 did not complete the posttest for various personal reasons (eg, illness, Lunar New Year gatherings, and concerns about social distancing during the late COVID-19 pandemic). This sample size met our calculated sample size for this intervention (with effect size of 0.5 [Cohen *d*]) in order to achieve a power of 80% at a significance level of 5% [19]. The retention rate was 82.6%, slightly higher than the expected 80%. Ultimately, 276 participants (LVB group, n=100; Taichi group, n=86; control group, n=90) who completed both the pre- and posttest were included in the final data analysis.

Table 1 presents the sociodemographic characteristics of the participants. Approximately 56% (156/276) of the participants were aged ≤70 years, and only 8% (22/276) were aged ≥80 years. The majority of the participants were female (229/276, 83%) and retired (217/276, 78.6%). Approximately 57.3% (158/276) had attained secondary education or higher. Approximately half (161/276, 58.3%) reported their financial status as average. The average BMI was 24 (SD 3.6) kg/m^2. The groups did not significantly differ with each other at the pretest with respect to all variables ($P>.05$).

Figure 1. Recruitment statistics.

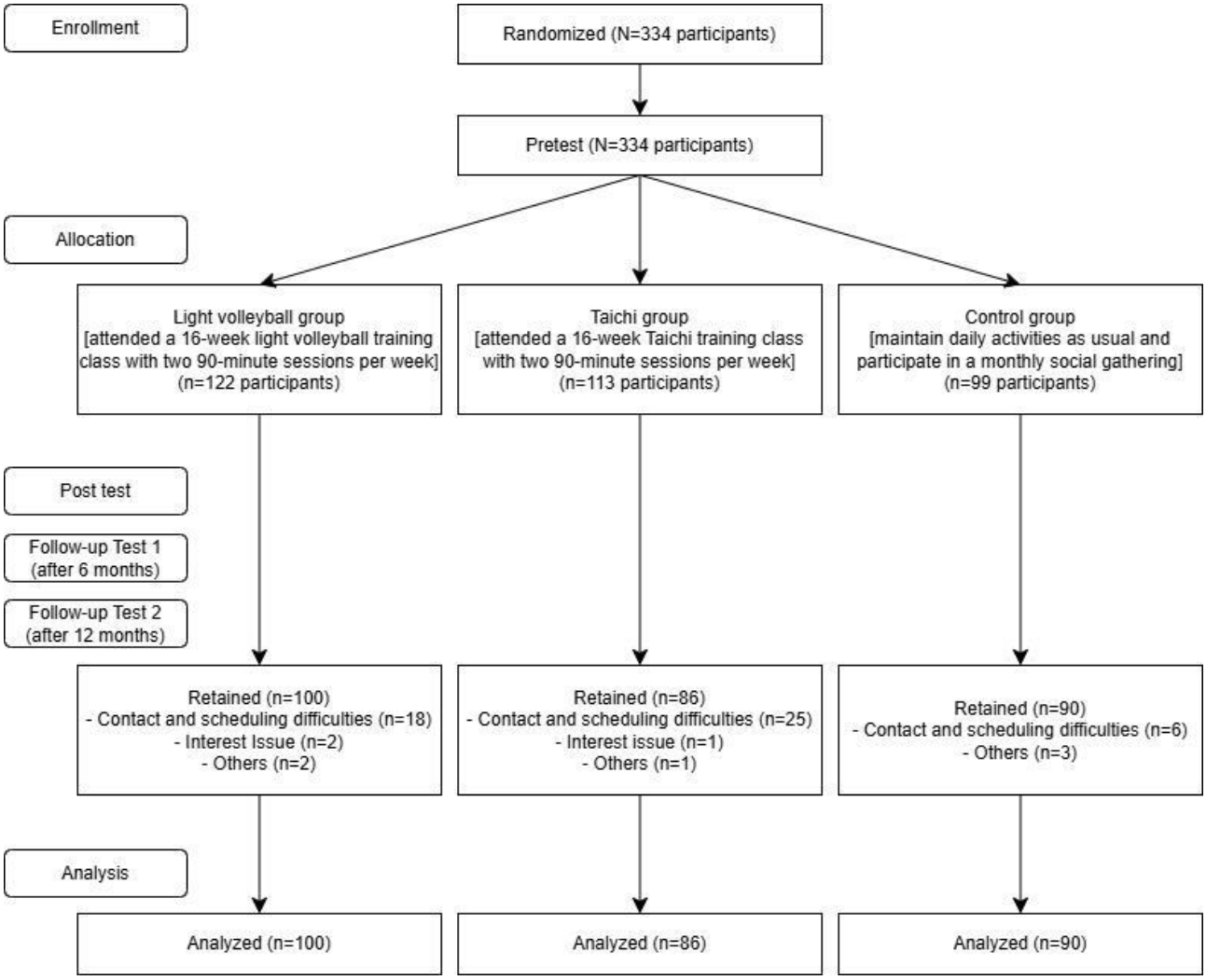


Table . Sociodemographic characteristics of participants.

Characteristics	LVB[a] (n=100)	TC[b] (n=86)	Control[c] (n=90)
Age (years), n (%)			
60 - 64	30 (30)	14 (16.3)	10 (11.1)
65 - 69	35 (35)	34 (39.5)	33 (36.7)
70 - 74	24 (24)	26 (30.2)	17 (18.9)
75 - 79	8 (8)	8 (9.3)	15 (16.7)
80 or above	3 (3)	4 (4.7)	15 (16.7)
Sex, n (%)			
Male	17 (17)	18 (20.9)	12 (13.3)
Female	83 (83)	68 (79.1)	78 (86.7)
Occupation, n (%)			
Full-time job	0 (0)	0 (0)	1 (1.1)
Housewife	16 (16)	11 (12.8)	21 (23.3)
Retired	79 (79)	72 (83.7)	66 (73.3)
Part-time job or others	5 (5)	3 (3.5)	2 (2.2)
Education level, n (%)			
No education	5 (5)	4 (4.7)	21 (23.3)
Primary education	35 (35)	23 (26.7)	30 (33.3)
Secondary education	40 (40)	41 (47.7)	29 (32.2)
Tertiary education	20 (20%)	18 (20.9%)	10 (11.1%)
Perceived financial status, n (%)			
Low	19 (19)	18 (20.9)	18 (20)
Below average	21 (21)	17 (19.8)	10 (11.1)
Average	53 (53)	50 (58.1)	58 (64.4)
Above average	3 (3)	1 (1.2)	3 (3.3)
Higher	4 (4)	0 (0)	1 (1.1)
House nature, n (%)			
Bought	46 (46)	33 (38.4)	32 (35.6)
Rent	54 (54)	53 (61.6)	58 (64.4)
BMI (kg/m^2), mean (SD)	24.1 (3.7)	23.5 (3.7)	24.2 (3.4)

[a]LVB: light volleyball.

[b]TC: Taichi.

[c]CG: control group.

Improvement in Physical Fitness

Table 2 presents the mean and SD values of physical measures for the groups at pre- and posttest, along with the results of analysis of covariance with repeated measures, with BMI controlled for. After the intervention, the LVB group achieved statistically significant improvements in lower limb muscular endurance (chair stand test, $F_{2,272}$= 7.2, P=.001, η^2=.05), agility (up-and-go test, $F_{2,272}$= 6.1, P=.003, η^2=.04), and dynamic balance (Biodex balance test, $F_{2,272}$= 9.4, P=.001, η^2=.07) compared with the other groups.

Pairwise comparisons revealed that both the LVB and Taichi groups exhibited significantly improved adjusted performance in lower limb muscular endurance (chair stand test, LVB group, mean 17.6, SD 5.0; Taichi group, mean 15.3, SD 4.9, P=.001), agility (up-and-go test, LVB group, mean 5.6, SD 1.3; Taichi group, mean 6.2, P=.001), and dynamic balance (Biodex balance test, LVB group, mean 0.5, SD 0.5; Taichi group, mean 0.7, SD 0.6, P=.001) compared with the control group (chair stand test, mean 15.3, SD 4.9; up-and-go test, mean 6.1, SD 1.2; Biodex balance test, mean 0.7, SD 0.5). Notably, the LVB group performed significantly better on all three measures (chair stand test, up-and-go test, and Biodex balance test) at posttest compared with the Taichi group. No significant group

differences were identified for the arm curl test ($F_{2,270}$=2.8, P=.1, partial η^2=.02), chair sit-and-reach test ($F_{2,270}$=0.06, P=.5, partial η^2=.005), and back scratch test ($F_{2,270}$=1.3, P=.3, partial η^2=.01), and 2-min step test ($F_{2,270}$=1.1, P=.3, partial η^2=.008).

Table . Mean and SD values for measures in groups at pre- and posttest.

Measures	LVB[a] (n=100), mean (SD)		TC[b] (n=86), mean (SD)		CG[c] (n=90), mean (SD)		*F* test (*df*)	Mean difference (95% CI)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest		LVB-TC	LVB-CG
Chair stand test (frequency)	16.4 (5.1)	17.6 (5.0)	15 (5.1)	15.3 (4.9)	15 (3.7)	15.3 (3.5)	7.2[d] (2, 272)	2.4[d] (0.8 to 3.9)	2.3[d] (0.8 to 3.9)
Arm curl (frequency)	15.3 (4.9)	16.0 (4.8)	14.4 (5.1)	14.6 (4.9)	14.1 (5.3)	14.2 (5.0)	2.8 (2, 270)	1.5 (−0.2 to 3.2)	1.9[e] (0.2 to 3.5)
Chair sit-and-reach test (cm)	7.3 (11.5)	7.8 (11.4)	5 (13.0)	5.2 (12.6)	5.1 (9.5)	4.8 (8.8)	0.7 (2, 272)	2.7 (−0.8 to 6.1)	3 (−0.4 to 6.4)
Back scratch (cm)	0.6 (9.3)	1.0 (9.5)	0.9 (8.9)	0.1 (8.7)	−2 (9.2)	−2.7 (9.4)	1.3 (2, 270)	0.9 (−1.9 to 3.6)	3.7[e] (0.1 to 6.3)
Up-and-go test (s)	6 (1.5)	5.6 (1.3)	6.2 (1.8)	6.2 (1.7)	6.5 (1.4)	6.1 (1.2)	6.1[d] (2, 272)	−.6[e] (−1.1 to −0.1)	−.7[d] (−1.2 to −0.2)
Step test (frequency)	92.1 (18.3)	94.6 (17.6)	89.4 (22.9)	90.0 (21.8)	85.3 (17.6)	85.9 (16.7)	1.3 (2, 272)	4.56 (−1.5 to 10.7)	8.7[e] (2.7 to 14.7)
Overall stability index (score)	0.7 (0.6)	0.5 (0.5)	0.7 (0.6)	0.7 (0.6)	0.8 (0.6)	0.7(0.5)	9.4[d] (2, 272)	−0.2[e] (−0.3 to −.01)	−0.2[d] (−0.4 to −0.04)

[a]LVB: light volleyball.

[b]TC: Taichi.

[c]CG: control group.

[d]P<.01.

[e]P<.05.

Discussion

Principal Findings

Given the continued aging of the population in Hong Kong and the limitations of our previous LVB pilot study, the current study aimed to investigate the effect of a 16-week LVB intervention program on physical health outcomes among older adults. Compared with the pilot study, the current study used an randomized controlled trial design and recruited 5.45 times more participants. The results indicated that the LVB intervention had a greater effect on improving lower body strength, agility, and dynamic balance in older adults when compared with both the Taichi intervention and the control group. Although significant improvements in lower body strength, agility, and dynamic balance were identified in the LVB intervention group, no significant changes in aerobic endurance and upper body strength were identified.

Improvement in Physical Health

The current study hypothesized that the LVB intervention would lead to greater improvements in lower body strength, agility, and dynamic balance relative to the control. This expectation is supported by our pilot study in 2018, which revealed significant improvements in these physical attributes for the LVB intervention group when compared with the control group [20].

We also hypothesized that the LVB intervention would improve lower body strength, agility, and dynamic balance relative to Taichi. First, LVB shares several similarities with traditional volleyball, such as involving a considerable amount of lower body movement, which helps enhance lower body strength [36]. Second, LVB was expected to result in more significant improvements in agility and dynamic balance compared with Taichi. This is because LVB requires players to use open skills, enabling them to react and adapt to dynamic, constantly changing environments. In contrast, Taichi primarily involves closed skills, which do not require such adaptive responses to environmental changes [37]. Although there may be differences, Sheppard and Young [38] defined agility as the ability to change speed or direction in response to a stimulus, such as an environmental change. A recent systematic review comparing the effects of open-skill physical activity and closed-skill physical activity on cognitive function reported that open-skill physical activity was more effective in enhancing cognitive

function than closed-skill exercises were [39]. Furthermore, Young et al [40] suggested that agility is strongly linked to cognitive functions such as decision-making ability and perception.

Additionally, a strong correlation has been observed between lower body strength and the ability to change direction (change-of-direction ability) [41]. Recent research has demonstrated a significant correlation between relative and absolute strength and agility, including change-of-direction ability and linear speed [42]. In the current study, we also identified a correlation between lower body strength and agility. These were the 2 physical attributes for which more pronounced improvements were identified in the LVB intervention group than in the Taichi group. In addition to lower body strength, balance plays a crucial role in maintaining good posture during acceleration, deceleration, and sudden changes in location or direction. Balance training has been found to be beneficial in improving the agility of volleyball players [43].

Although the participants in the LVB group exhibited greater improvements in lower body strength, agility, and dynamic balance, an unexpected finding was the minimal differences in upper body strength and aerobic endurance, 2 physical attributes that were initially hypothesized to exhibit significant improvement in the LVB group compared with the Taichi group. For upper body strength, LVB involves more frequent and vigorous arm and shoulder movements relative to Taichi, such as spiking and blocking, which were expected to result in greater upper body strength improvements compared with Taichi [44]. Similarly, the higher energy demand in volleyball was assumed to lead to greater improvements in aerobic endurance in the LVB group compared with Taichi [45]. However, despite these expectations, no significant differences in improvements were observed between the LVB and Taichi groups for these two physical attributes. Nonetheless, both the LVB and Taichi programs were demonstrated to improve upper body strength and aerobic endurance in older adults, although LVB did not significantly outperform Taichi as was hypothesized. Previous studies have indicated that both LVB and Taichi can be beneficial for improving upper body muscle strength in older adults [19,44]. Regarding aerobic capacity, although we assumed that volleyball would require higher aerobic endurance,Taichi's breathing techniques were found to be beneficial for improving aerobic capacity in older adults, providing benefits similar to those achieved through LVB [46].

Limitations

Although this study addressed the limitations of the previous pilot study, it still has several limitations of its own. First, these are only preliminary findings, and the scheduled follow-up evaluation in 6 or 12 months will help determine whether these advantages are seen or sustained over the longer term after the intervention. Second, the gender representation was disproportionate; the number of female participants was higher because recruitment was conducted through local elderly centers. These centers in Hong Kong have predominantly female memberships, with older women showing a greater tendency to participate in activities [47,48]. This imbalance in gender representation limits the generalizability of the findings. Furthermore, the study was conducted during the COVID-19 pandemic and the proposed intervention period was postponed from 2020 - 2022 to 2021 - 2023, which may have introduced variables that are not typically present in nonpandemic conditions. These factors could have affected the results of the interventions.

Conclusions

The current study examined the effects of a 16-week LVB intervention program on the physical health of older adults in Hong Kong. The results indicated that the participants in the LVB group experienced significant improvements in lower body strength, agility, and dynamic balance compared with both the Taichi active control group and the control group. This study builds on the previous pilot by adopting an randomized controlled trial design, incorporating dynamic balance as a fitness component, increasing the sample size, and collecting data from 7 local elderly centers instead of just one.

Future studies should address the following: first, studies should aim for more balanced gender representation and include follow-up tests to monitor the long-term maintenance of physical improvements. Numerous studies have highlighted the positive effects of physical activity on aspects such as cognition and psychology in older adults, whereas the current study focused solely on physical attributes. Future research could explore the effects of LVB on cognitive and psychological outcomes to provide a more comprehensive understanding of its benefits. Furthermore, future studies should consider qualitative approaches to understand participants' experiences and assess how LVB intervention influences various health dimensions. Given the potential limitations due to the COVID-19 pandemic, future studies may benefit from replication under more stable conditions, free from social distancing policies, to determine whether similar results can be observed. In conclusion, this large-scale study provides strong evidence supporting the physical health benefits of LVB for older adults. The results from our future qualitative studies and follow-up measures will further inform researchers and practitioners about the acceptability and appropriateness of LVB interventions for older adults.

Conflicts of Interest

None declared.

Checklist 1
CONSORT Checklist
[PDF File, 227 KB - aging_v8i1e62886_app1.pdf]

References

1. Persons with disabilities and chronic diseases. Census and Statistics Department, Hong Kong Special Administrative Region. 2020. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/C0000055/att/B11301632021XXXXB0100.pdf [accessed 2025-04-08]
2. Synopsis of healthcare financing studies projection of Hong Kong's healthcare expenditure. Food and Health Bureau, Hong Kong Special Administrative Region Government. 2008. URL: https://www.healthbureau.gov.hk/beStrong/files/consultation/projecthealthexp_eng.pdf [accessed 2025-04-08]
3. Moreno-Agostino D, Daskalopoulou C, Wu YT, et al. The impact of physical activity on healthy ageing trajectories: evidence from eight cohort studies. Int J Behav Nutr Phys Act 2020 Jul 16;17(1):92. [doi: 10.1186/s12966-020-00995-8] [Medline: 32677960]
4. Territory-wide physical fitness survey for the community. Sports Commission of Hong Kong Special Administrative Region Government. 2024. URL: https://www.legco.gov.hk/yr2023/english/panels/ha/papers/ha20230612cb2-539-5-e.pdf [accessed 2025-04-08]
5. Healthy exercise for all campaign - fitness programmes for the elderly. Leisure and Cultural Services Department, Hong Kong Special Administrative Region Government. 2023. URL: https://www.lcsd.gov.hk/en/healthy/fitness/elderly.html [accessed 2025-04-08]
6. Elderly commission: report on healthy ageing executive summary. The Government of Hong Kong Special Administrative Region fo the People's Republic of China. 2022. URL: https://www.elderlycommission.gov.hk/en/library/Ex-sum.htm [accessed 2025-04-08]
7. Chief Secretary for Administration's Office. Population policy strategies and initiatives. 2016. URL: https://gia.info.gov.hk/general/201501/15/P201501150880_0880_140608.pdf [accessed 2025-04-08]
8. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
9. World Health Organization. Active ageing: a policy framework (no. WHO/NMH/NPH/02.8). 2004. URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHO-Active-Ageing-Framework.pdf [accessed 2025-04-08]
10. Leisure and Cultural Services Department, Hong Kong Special Administrative Region Government. Use of public open space. 2016. URL: https://www.legco.gov.hk/yr15-16/english/panels/ha/papers/ha20160614cb2-1680-3-e.pdf [accessed 2025-04-08]
11. Sherrington C, Michaleff ZA, Fairhall N, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2017 Dec;51(24):1750-1758. [doi: 10.1136/bjsports-2016-096547] [Medline: 27707740]
12. Sun M, Min L, Xu N, Huang L, Li X. The effect of exercise intervention on reducing the fall risk in older adults: a meta-analysis of randomized controlled trials. Int J Environ Res Public Health 2021 Nov 29;18(23):12562. [doi: 10.3390/ijerph182312562] [Medline: 34886293]
13. Chase JAD, Phillips LJ, Brown M. Physical activity intervention effects on physical function among community-dwelling older adults: a systematic review and meta-analysis. J Aging Phys Act 2017 Jan;25(1):149-170. [doi: 10.1123/japa.2016-0040] [Medline: 27620705]
14. Taylor J, Walsh S, Kwok W, et al. A scoping review of physical activity interventions for older adults. Int J Behav Nutr Phys Act 2021 Jun 30;18(1):82. [doi: 10.1186/s12966-021-01140-9] [Medline: 34193157]
15. Pinheiro MB, Reis AHS, Baldwin JN, et al. Quantity and quality are increasing but there's room for improvement: A scoping review of physical activity intervention trials. Braz J Phys Ther 2024;28(2):101051. [doi: 10.1016/j.bjpt.2024.101051] [Medline: 38574557]
16. Wong MYC, Ou KL, Chung PK, Chui KYK, Zhang CQ. The relationship between physical activity, physical health, and mental health among older Chinese adults: A scoping review. Front Public Health 2022;10:914548. [doi: 10.3389/fpubh.2022.914548] [Medline: 36684983]
17. Franco MR, Tong A, Howard K, et al. Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. Br J Sports Med 2015 Oct;49(19):1268-1276. [doi: 10.1136/bjsports-2014-094015] [Medline: 25586911]
18. Van Dyck D, Mertens L, Cardon G, De Cocker K, De Bourdeaudhuij I. Opinions toward physical activity, sedentary behavior, and interventions to stimulate active living during early retirement: a qualitative study in recently retired adults. J Aging Phys Act 2017 Apr;25(2):277-286. [doi: 10.1123/japa.2015-0295] [Medline: 27759483]
19. Leung KM, Chung PK, Chan AWK, et al. Promoting healthy ageing through light volleyball intervention in Hong Kong: study protocol for a randomised controlled trial. BMC Sports Sci Med Rehabil 2020;12(1):6. [doi: 10.1186/s13102-019-0151-7] [Medline: 32002186]
20. Leung KM, Chung PK, Hagger MS. The effects of light volleyball intervention programme in improving selected physical and psychological attributes of older adults in Hong Kong. Int J Sport Exerc Psychol 2020 Jan 2;18(1):1-12. [doi: 10.1080/1612197X.2018.1462231]

21. Chan AWK, Lee A, Lee DTF, et al. The sustaining effects of Tai chi Qigong on physiological health for COPD patients: a randomized controlled trial. Complement Ther Med 2013 Dec;21(6):585-594. [doi: 10.1016/j.ctim.2013.09.008] [Medline: 24280465]
22. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JGZ. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56. [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
23. Burke SM, Carron AV, Eys MA, Ntoumanis N, Estabrooks PA. Group versus individual approach? A meta-analysis of the effectiveness of interventions to promote physical activity. Int Rev Sport Exerc Psychol 2006 Feb;2(1):13-29. [doi: 10.53841/bpssepr.2006.2.1.13]
24. Healthy exercise for all campaign – physical fitness test for the community. Sports Commission of Hong Kong Special Administrative Region Government. 2012. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/FA100265/att/B71302FA2013XXXXB0100.pdf [accessed 2025-04-08]
25. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166. [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
26. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8(1):18. [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
27. Shekelle PG, Maglione MA, Mojica W, et al. Exercise programs for older adults: a systematic review and meta-analysis. RAND. 2003 URL: https://www.rand.org/pubs/reprints/RP1257.html [accessed 2025-05-08]
28. Chu LW, Pei CKW, Ho MH. Validation of the abbreviated mental test (Hong Kong version) in the elderly medical patient. Hong Kong Med J 1995:207-211.
29. Schoene D, Wu S, Mikolaizak AS, Wu S‐, et al. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta‐analysis. J American Geriatrics Society 2013 Feb;61(2):202-208 [FREE Full text] [doi: 10.1111/jgs.12106]
30. Rikli RE, Jones CJ. Senior fitness test manual. In: Human Kinetics 2013.
31. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999 Jan;7(2):129-161. [doi: 10.1123/japa.7.2.129]
32. Parraca JA, Olivares PR, Carbonell-Baeza A, Aparicio VA, Adsuar JC, Gusi N. Test-retest reliability of Biodex Balance SD on physically active old people. JHSE 2011 Jun 1;6(2):444-451 [FREE Full text] [doi: 10.4100/jhse.2011.62.25]
33. Finn JA, Alvarez MM, Jett RE, Axtell RS, Kemler DS. Stability performance assessment among subjects of disparate balancing abilities. Med Sci Sports Exerc 1999 May;31(Supplement):S252. [doi: 10.1097/00005768-199905001-01209]
34. Riebe D, Blissmer BJ, Greaney ML, Garber CE, Lees FD, Clark PG. The relationship between obesity, physical activity, and physical function in older adults. J Aging Health 2009 Dec;21(8):1159-1178. [doi: 10.1177/0898264309350076] [Medline: 19897781]
35. Cohen J. Statistical power analysis for the behavioral sciences. Comput Environ Urban Syst 1990 Jan;14(1):71. [doi: 10.1016/0198-9715(90)90050-4]
36. Tramel W, Lockie RG, Lindsay KG, Dawes JJ. Associations between absolute and relative lower body strength to measures of power and change of direction speed in division II female volleyball players. Sports (Basel) 2019 Jul 1;7(7):160. [doi: 10.3390/sports7070160] [Medline: 31266193]
37. Koch P, Krenn B. Executive functions in elite athletes – comparing open-skill and closed-skill sports and considering the role of athletes' past involvement in both sport categories. Psychol Sport Exerc 2021 Jul;55:101925. [doi: 10.1016/j.psychsport.2021.101925]
38. Sheppard JM, Young WB. Agility literature review: classifications, training and testing. J Sports Sci 2006 Sep;24(9):919-932. [doi: 10.1080/02640410500457109] [Medline: 16882626]
39. Gu Q, Zou L, Loprinzi PD, Quan M, Huang T. Effects of open versus closed skill exercise on cognitive function: a systematic review. Front Psychol 2019;10:1707. [doi: 10.3389/fpsyg.2019.01707] [Medline: 31507472]
40. Young WB, Dawson B, Henry GJ. Agility and change-of-direction speed are independent skills: implications for training for agility in invasion sports. Int J Sports Sci Coach 2015 Feb;10(1):159-169. [doi: 10.1260/1747-9541.10.1.159]
41. Sheppard JM, Dawes JJ, Jeffreys I, Spiteri T, Nimphius S. Broadening the view of agility: a scientific review of the literature. Journal of Australian Strength & Conditioning 2014;22(3):6-25 [FREE Full text]
42. Andersen E, Lockie RG, Dawes JJ. Relationship of absolute and relative lower-body strength to predictors of athletic performance in collegiate women soccer players. Sports (Basel) 2018 Sep 29;6(4):106. [doi: 10.3390/sports6040106] [Medline: 30274252]
43. Gadre HV, Lele DC, Deo M, Mathur C. Effect of dynamic balance training on agility in adolescent volleyball players. JCDR 2019 Jan 1. [doi: 10.7860/JCDR/2019/42282.13310]
44. Hedrick A. Training for high level performance in women's collegiate volleyball: part I training requirements. J Strength Cond 2007;29(6):50. [doi: 10.1519/1533-4295(2007)29[50:TFHLPI]2.0.CO;2]
45. Wehner C, Blank C, Arvandi M, Wehner C, Schobersberger W. Effect of tai chi on muscle strength, physical endurance, postural balance and flexibility: a systematic review and meta-analysis. BMJ Open Sport Exerc Med 2021;7(1):e000817. [doi: 10.1136/bmjsem-2020-000817] [Medline: 33614126]

46. Zou L, Loprinzi PD, Yu JJ, et al. Superior effects of modified Chen-style tai chi versus 24-style tai chi on cognitive function, fitness, and balance performance in adults over 55. Brain Sci 2019 May 4;9(5):102. [doi: 10.3390/brainsci9050102] [Medline: 31060221]
47. Caritas hong kong – servies for the elderly. Service Communications. 2007. URL: https://www.caritasse.org.hk/zh/wp-content/uploads/2020/07/%E5%AE%89%E8%80%81%E6%9C%8D%E5%8B%99%E9%80%9A%E8%A8%8A_%E7%AC%AC11%E6%9C%9F.pdf [accessed 2025-04-08]
48. St. James Settlement. The Voice Issue No 350 The Voice 2006 [Website in Chinese]. URL: https://www.thevoice.org.hk/v0350/00002.htm [accessed 2025-05-08]

Abbreviations

AMT : Abbreviated Mental Test
CONSORT: Consolidated Standards of Reporting Trials
LVB: light volleyball
TUG: Timed-Up-and-Go
WHO: World Health Organization

Edited by J Job; submitted 03.06.24; peer-reviewed by C Sherrington, R Strauss; revised version received 22.01.25; accepted 29.01.25; published 28.05.25.

Please cite as:
Leung KM, Shi Y
Impact of a Light Volleyball Intervention Program on Improving Physical Attributes of Older Adults in Hong Kong: Preliminary Study of a Randomized Controlled Trial
JMIR Aging 2025;8:e62886
URL: https://aging.jmir.org/2025/1/e62886
doi:10.2196/62886

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

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

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

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

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

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

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

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

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

Abstract

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

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

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

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

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

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

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

KEYWORDS

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

Introduction

Background

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

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

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

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

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

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

Objectives

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

Methods

Study Design

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

Ethical Considerations

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

Participants

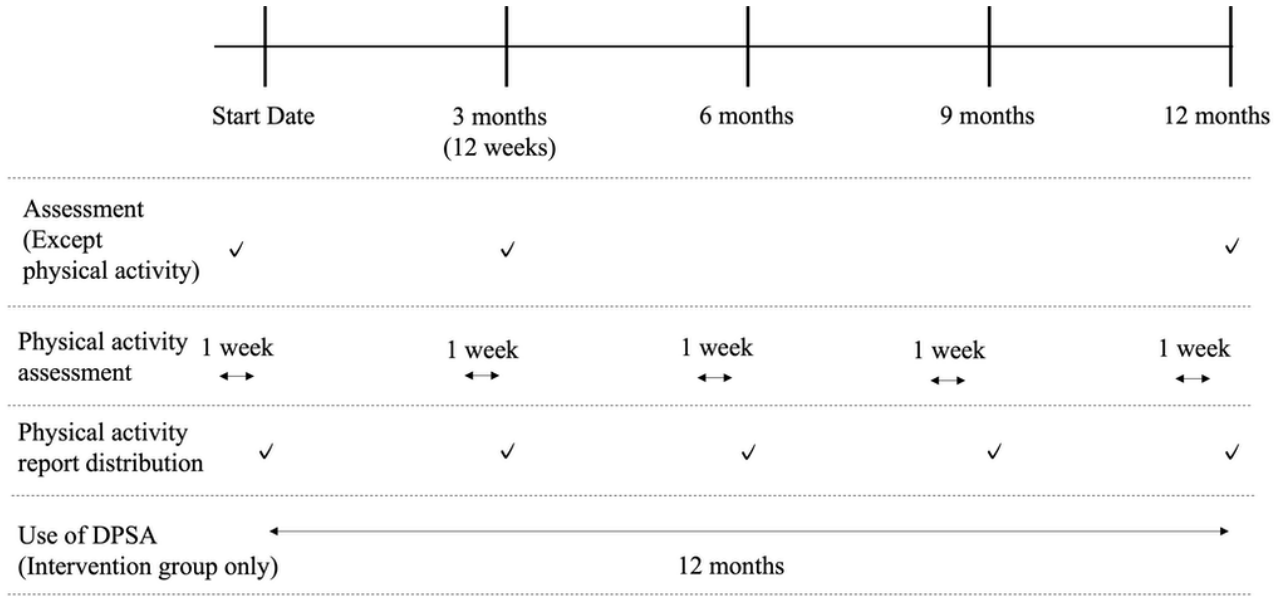
The study included Fujisawa City residents aged 65 years or older. In Japan, older adults are generally defined as persons aged 65 years or older [40]. We recruited participants for two 3-month programs designed to increase physical activity [30]. Participants from 2 different areas within the municipality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

Intervention

Program

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

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



Digital Peer-Supported App

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

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

Characteristics of Research Participants

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

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

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

DPSA Feasibility

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

Outcome Measure

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

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

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

Statistical Analysis

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

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

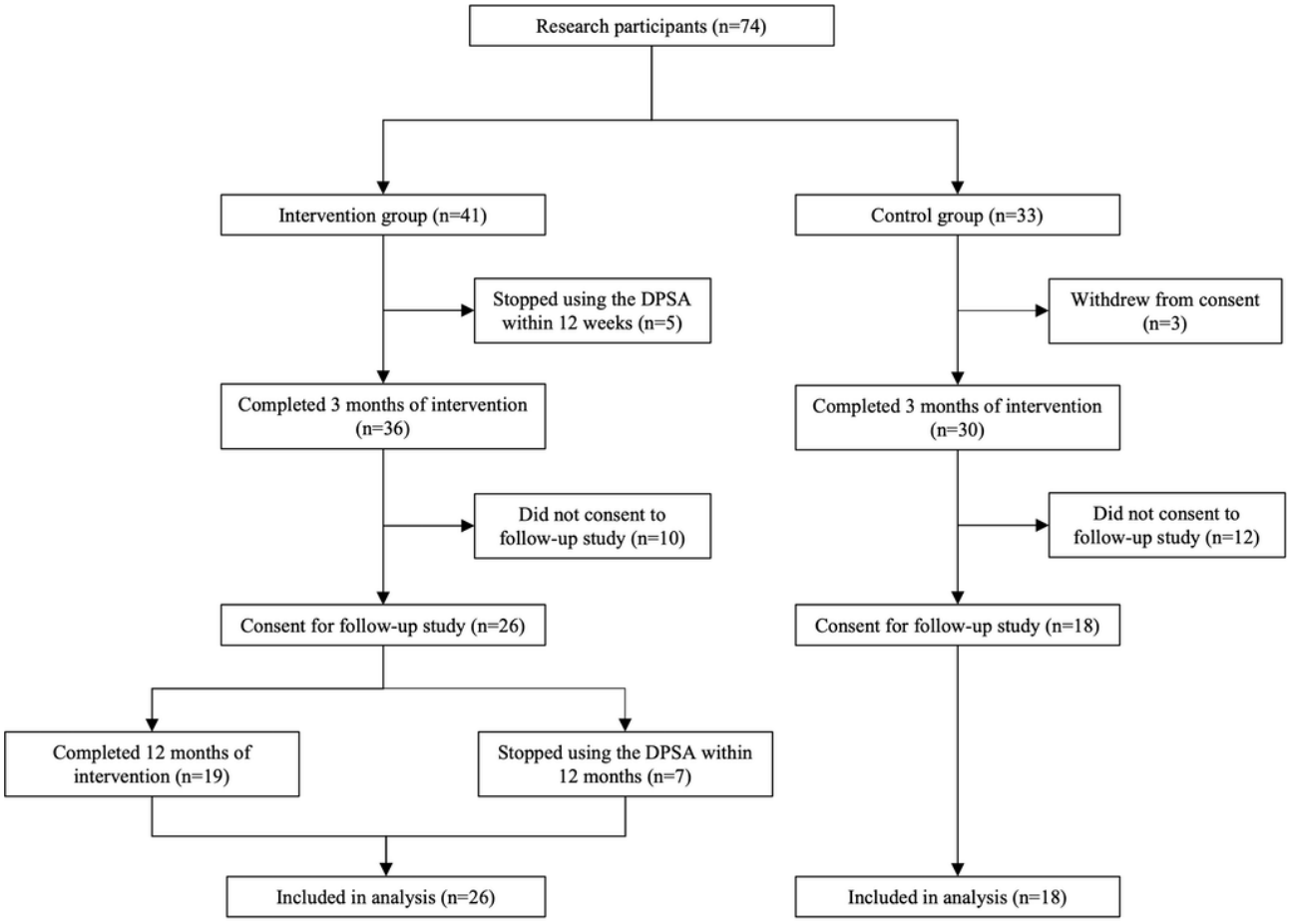
Data were analyzed using IBM SPSS Statistics 29 (IBM). The statistical significance level was set to 5%.

Results

Participants

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

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



Participant Characteristics

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

Table . Participant characteristics.

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

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

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

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

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

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

[e]LPA: light-intensity physical activity.

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

[g]SB: sedentary behavior.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[c]LPA: light-intensity physical activity.

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

[e]SB: sedentary behavior.

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

Discussion

Principal Results

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

Comparison With Previous Studies

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

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

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

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

Limitations

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

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

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

Fourth, the generalizability of this study is limited owing to potential selection bias. Participants in the intervention group were not only motivated to increase physical activity but also familiar with using the app. The DPSA was not adaptable to all participants, as it required a certain level of information technology literacy and cognitive function. Social, cultural, and economic factors (eg, older age, privacy concerns, and low income) may influence preference and feasibility with smartphone apps [81-83]. Therefore, the use of DPSA may not be widely accepted.

Conclusions

This study assessed the 12-month feasibility of using the DPSA and measured preliminary estimates of its effects on physical activity, physical function, and self-efficacy for exercise. The 12-month retention rate for participants in the DPSA intervention group was 73% (19/26), and that for 41 participants including those who decided not to participate in the follow-up study was 46% (19/41). The DPSA adherence rate was 85.9%. Only a limited number of older adults had long-term access to the DPSA. Preliminary estimates suggest that DPSA use may improve step count, lower extremity function, and self-efficacy for exercise. However, various biases were introduced, preventing the demonstration of clear intervention effects. There

is a need to identify ways in which more older adults can use DPSAs for extended periods of time; RCTs should be conducted to ascertain the long-term effects of DPSAs on physical activity and function in older adults.

Acknowledgments

This study was supported by a Doctoral Student Grant-in-Aid Program from the Ushioda Memorial Fund. The Fujisawa City government, A10 Lab Inc., and Keio University independently funded their respective parts of the research. The authors thank the staff of the Elderly Support Division of the Fujisawa City government (Shinichi Sakae, Mayumi Taguchi, Mari Sakata, and Shizuka Kobayashi) for recruiting participants and managing the project, Ken Takahashi of the Fujisawa City Health and Medical Foundation for his cooperation as a health fitness instructor, and Oguma laboratory staff for their help in collecting the data.

Authors' Contributions

KT, YO, SY, and MS were responsible for designing this study. KT, YO, SY, MS, MN, ND, AH, and TM were responsible for the data collection. KT and YO were responsible for analyzing and interpreting the data. All authors have edited, reviewed, and approved the final manuscript.

Conflicts of Interest

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

Multimedia Appendix 1
Details of the digital peer-supported app's functionality.
[DOCX File, 70532 KB - aging_v8i1e66610_app1.docx]

References

1. Zhang D. Perceived neighborhood conditions, psychosocial factors, and sleep problems among urban and rural older adults in China. J Aging Health 2024 Jun;36(5-6):337-349. [doi: 10.1177/08982643231159709] [Medline: 37395560]
2. Lutz W, Sanderson W, Scherbov S. Doubling of world population unlikely. Nature New Biol 1997 Jun 19;387(6635):803-805. [doi: 10.1038/42935] [Medline: 9194559]
3. Lang PO, Govind S, Aspinall R. Reversing T cell immunosenescence: why, who, and how. Age (Dordr) 2013 Jun;35(3):609-620. [doi: 10.1007/s11357-012-9393-y] [Medline: 22367580]
4. Bouaziz W, Vogel T, Schmitt E, Kaltenbach G, Geny B, Lang PO. Health benefits of aerobic training programs in adults aged 70 and over: a systematic review. Arch Gerontol Geriatr 2017;69:110-127. [doi: 10.1016/j.archger.2016.10.012] [Medline: 27912156]
5. White paper on the aging society. Cabinet Office Japan. 2023. URL: https://www8.cao.go.jp/kourei/whitepaper/w-2024/zenbun/06pdf_index.html [accessed 2024-09-17]
6. Dementia. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/dementia#:~:text=Studies%20show%20that%20people%20can,cholesterol%20and%20blood%20sugar [accessed 2024-05-17]
7. Katzmarzyk PT, Friedenreich C, Shiroma EJ, Lee IM. Physical inactivity and non-communicable disease burden in low-income, middle-income and high-income countries. Br J Sports Med 2022 Jan;56(2):101-106. [doi: 10.1136/bjsports-2020-103640] [Medline: 33782046]
8. Manning KM, Hall KS, Sloane R, et al. Longitudinal analysis of physical function in older adults: the effects of physical inactivity and exercise training. Aging Cell 2024 Jan;23(1):e13987. [doi: 10.1111/acel.13987] [Medline: 37681737]
9. Wang H, Huang WY, Zhao Y. Efficacy of exercise on muscle function and physical performance in older adults with sarcopenia: an updated systematic review and meta-analysis. Int J Environ Res Public Health 2022 Jul 5;19(13):8212. [doi: 10.3390/ijerph19138212] [Medline: 35805870]
10. Lee IM, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229. [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
11. Promotion of physical activity and exercise 2023. Ministry of Health, Labour and Welfare. 2024 Jan. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/kenkou/undou/index.html [accessed 2025-02-12]
12. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]
13. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983 Dec;31(12):721-727. [doi: 10.1111/j.1532-5415.1983.tb03391.x] [Medline: 6418786]
14. Rejeski WJ, Ip EH, Marsh AP, Miller ME, Farmer DF. Measuring disability in older adults: the International Classification System of Functioning, Disability and Health (ICF) framework. Geriatr Gerontol Int 2008 Mar;8(1):48-54. [doi: 10.1111/j.1447-0594.2008.00446.x] [Medline: 18713189]

15. Fukuoka Y, Haskell W, Lin F, Vittinghoff E. Short- and long-term effects of a mobile phone app in conjunction with brief in-person counseling on physical activity among physically inactive women: the mPED randomized clinical trial. JAMA Netw Open 2019 May 3;2(5):e194281. [doi: 10.1001/jamanetworkopen.2019.4281] [Medline: 31125101]
16. Laranjo L, Ding D, Heleno B, et al. Do smartphone applications and activity trackers increase physical activity in adults? Systematic review, meta-analysis and metaregression. Br J Sports Med 2021 Apr;55(8):422-432. [doi: 10.1136/bjsports-2020-102892] [Medline: 33355160]
17. Moss RJ, Süle A, Kohl S. eHealth and mHealth. Eur J Hosp Pharm 2019 Jan;26(1):57-58. [doi: 10.1136/ejhpharm-2018-001819] [Medline: 31157099]
18. Elavsky S, Knapova L, Klocek A, Smahel D. Mobile health interventions for physical activity, sedentary behavior, and sleep in adults aged 50 years and older: a systematic literature review. J Aging Phys Act 2019 Aug 1;27(4):565-593. [doi: 10.1123/japa.2017-0410] [Medline: 30507266]
19. Kwan RYC, Salihu D, Lee PH, et al. The effect of e-health interventions promoting physical activity in older people: a systematic review and meta-analysis. Eur Rev Aging Phys Act 2020;17:7. [doi: 10.1186/s11556-020-00239-5] [Medline: 32336996]
20. D'Amore C, Reid JC, Chan M, et al. Interventions including smart technology compared with face-to-face physical activity interventions in older adults: systematic review and meta-analysis. J Med Internet Res 2022 Oct 31;24(10):e36134. [doi: 10.2196/36134] [Medline: 36315229]
21. Duan Y, Shang B, Liang W, Du G, Yang M, Rhodes RE. Effects of eHealth-based multiple health behavior change interventions on physical activity, healthy diet, and weight in people with noncommunicable diseases: systematic review and meta-analysis. J Med Internet Res 2021 Feb 22;23(2):e23786. [doi: 10.2196/23786] [Medline: 33616534]
22. Morey MC, Snyder DC, Sloane R, et al. Effects of home-based diet and exercise on functional outcomes among older, overweight long-term cancer survivors: RENEW: a randomized controlled trial. JAMA 2009 May 13;301(18):1883-1891. [doi: 10.1001/jama.2009.643] [Medline: 19436015]
23. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: International Bibliography of the Social Sciences (IBSS); 1986.
24. McCullagh P, Weiss MR, Ross D. Modeling considerations in motor skill acquisition and performance: an integrated approach. Exerc Sport Sci Rev 1989;17:475-513. [Medline: 2676555]
25. Posavac EJ, Kattapong KR, Dew DE Jr. Peer-based interventions to influence health-related behaviors and attitudes: a meta-analysis. Psychol Rep 1999 Dec;85(3 Pt 2):1179-1194. [doi: 10.2466/pr0.1999.85.3f.1179] [Medline: 10710973]
26. Ashford D, Bennett SJ, Davids K. Observational modeling effects for movement dynamics and movement outcome measures across differing task constraints: a meta-analysis. J Mot Behav 2006 May;38(3):185-205. [doi: 10.3200/JMBR.38.3.185-205] [Medline: 16709559]
27. McAuley E, Blissmer B. Self-efficacy determinants and consequences of physical activity. Exerc Sport Sci Rev 2000 Apr;28(2):85-88. [Medline: 10902091]
28. Ginis KAM, Nigg CR, Smith AL. Peer-delivered physical activity interventions: an overlooked opportunity for physical activity promotion. Transl Behav Med 2013 Dec;3(4):434-443. [doi: 10.1007/s13142-013-0215-2] [Medline: 24294332]
29. Webel AR, Okonsky J, Trompeta J, Holzemer WL. A systematic review of the effectiveness of peer-based interventions on health-related behaviors in adults. Am J Public Health 2010 Feb;100(2):247-253. [doi: 10.2105/AJPH.2008.149419] [Medline: 20019321]
30. Tabira K, Oguma Y, Yoshihara S, et al. Digital peer-supported app intervention to promote physical activity among community-dwelling older adults: nonrandomized controlled trial. JMIR Aging 2024 May 30;7:e56184. [doi: 10.2196/56184] [Medline: 38814686]
31. van der Meulen K, Granizo L, Del Barrio C. Emotional peer support interventions for students with SEND: a systematic review. Front Psychol 2021;12:797913. [doi: 10.3389/fpsyg.2021.797913] [Medline: 35027905]
32. Burton E, Farrier K, Hill KD, Codde J, Airey P, Hill AM. Effectiveness of peers in delivering programs or motivating older people to increase their participation in physical activity: systematic review and meta-analysis. J Sports Sci 2018 Mar;36(6):666-678. [doi: 10.1080/02640414.2017.1329549] [Medline: 28535358]
33. Iliffe S, Kendrick D, Morris R, et al. Multi-centre cluster randomised trial comparing a community group exercise programme with home based exercise with usual care for people aged 65 and over in primary care: protocol of the ProAct 65+ trial. Trials 2010 Jan 18;11:6. [doi: 10.1186/1745-6215-11-6] [Medline: 20082696]
34. McGarrigle L, Todd C. Promotion of physical activity in older people using mHealth and eHealth technologies: rapid review of reviews. J Med Internet Res 2020 Dec 29;22(12):e22201. [doi: 10.2196/22201] [Medline: 33372894]
35. Royse LA, Baker BS, Warne-Griggs MD, et al. "It's not time for us to sit down yet": how group exercise programs can motivate physical activity and overcome barriers in inactive older adults. Int J Qual Stud Health Well-being 2023 Dec;18(1):2216034. [doi: 10.1080/17482631.2023.2216034] [Medline: 37220010]
36. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity—a systematic review of longitudinal studies. BMC Public Health 2013 Sep 8;13:813. [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]

37. Jonkman NH, van Schooten KS, Maier AB, Pijnappels M. eHealth interventions to promote objectively measured physical activity in community-dwelling older people. Maturitas 2018 Jul;113:32-39. [doi: 10.1016/j.maturitas.2018.04.010] [Medline: 29903646]
38. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e14343. [doi: 10.2196/14343] [Medline: 31778121]
39. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699. [doi: 10.2196/26699] [Medline: 33811021]
40. Stecher C, Pfisterer B, Harden SM, et al. Assessing the pragmatic nature of mobile health interventions promoting physical activity: systematic review and meta-analysis. JMIR Mhealth Uhealth 2023 May 4;11:e43162. [doi: 10.2196/43162] [Medline: 37140972]
41. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing 2012 Jan;41(1):16-23. [doi: 10.1093/ageing/afr103]
42. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS One 2016;11(12):e0168614. [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
43. Fujisawa City future population estimates. Fujisawa City. 2023 Apr. URL: https://www.city.fujisawa.kanagawa.jp/documents/30105/fujisawasi-syourai-jinnkou-suikei-2022.pdf [accessed 2025-02-12]
44. Landry R, Traore N, Godin B. An econometric analysis of the effect of collaboration on academic research productivity. High Educ 1996 Oct;32(3):283-301. [doi: 10.1007/BF00138868]
45. Dooley L, Kirk D. University - industry collaboration. Eur J Innov Manage 2007 Aug 7;10(3):316-332. [doi: 10.1108/14601060710776734]
46. Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
47. Bredin SSD, Gledhill N, Jamnik VK, Warburton DER. PAR-Q+ and ePARmed-X+: new risk stratification and physical activity clearance strategy for physicians and patients alike. Can Fam Physician 2013 Mar;59(3):273-277. [Medline: 23486800]
48. Active guide: Japanese official physical activity guidelines for health promotion brochure in English. Ministry of Health, Labour and Welfare. 2013. URL: https://www.nibn.go.jp/eiken/programs/pdf/active2013-e.pdf [accessed 2025-12-02]
49. Fujisawa +10 exercise. Fujisawa City Health and Medical Center. 2014. URL: http://iryo.city.fujisawa.kanagawa.jp/health/health01_01.htm#02 [accessed 2025-01-17]
50. Osawa Y, Saito Y, Tsunekawa N, Manabe T, Oguma Y. Exercise workload of the "Fujisawa +10 exercise" program in older women. J Exer Physiol Online 2015;18:79-85.
51. Minchalle Web. A10Lab. Inc. 2023. URL: https://www.mhlw.go.jp/content/10904750/000999445.pdfhttps://minchalle.com/ [accessed 2025-02-12]
52. Ministry of Health, Labour and Welfare. National health and nutrition survey. 2017. URL: https://www.mhlw.go.jp/content/10904750/000351576.pdf [accessed 2025-02-12]
53. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166. [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]
54. Picorelli AMA, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156. [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
55. Ohkawara K, Oshima Y, Hikihara Y, Ishikawa-Takata K, Tabata I, Tanaka S. Real-time estimation of daily physical activity intensity by a triaxial accelerometer and a gravity-removal classification algorithm. Br J Nutr 2011 Jun;105(11):1681-1691. [doi: 10.1017/S0007114510005441] [Medline: 21262061]
56. Park J, Ishikawa-Takata K, Tanaka S, Bessyo K, Tanaka S, Kimura T. Accuracy of estimating step counts and intensity using accelerometers in older people with or without assistive devices. J Aging Phys Act 2017 Jan;25(1):41-50. [doi: 10.1123/japa.2015-0201] [Medline: 27180730]
57. Jefferis BJ, Sartini C, Ash S, et al. Trajectories of objectively measured physical activity in free-living older men. Med Sci Sports Exerc 2015 Feb;47(2):343-349. [doi: 10.1249/MSS.0000000000000410] [Medline: 24988411]
58. Savas S, Kilavuz A, Kayhan Koçak F, Cavdar S. Comparison of grip strength measurements by widely used three dynamometers in outpatients aged 60 years and over. J Clin Med 2023 Jun 25;12(13):4260. [doi: 10.3390/jcm12134260] [Medline: 37445293]
59. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
60. Oka K. Stages of change for exercise behavior and self-efficacy for exercise among middle-aged adults. Nihon Koshu Eisei Zasshi 2003 Mar;50(3):208-215. [Medline: 12704833]

61. Li Y, Baron J. Linear mixed-effects models in analyzing repeated-measures data. In: Li Y, Baron J, editors. Behavioral Research Data Analysis With R: Springer; 2012:177-204. [doi: 10.1007/978-1-4614-1238-0_10]
62. Raudenbush SW, Xiao-Feng L. Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. Psychol Methods 2001 Dec;6(4):387-401. [doi: 10.1037/1082-989X.6.4.387] [Medline: 11778679]
63. Muth C, Bales KL, Hinde K, Maninger N, Mendoza SP, Ferrer E. Alternative models for small samples in psychological research: applying linear mixed effects models and generalized estimating equations to repeated measures data. Educ Psychol Meas 2016 Feb;76(1):64-87. [doi: 10.1177/0013164415580432] [Medline: 29795857]
64. Wiley RW, Rapp B. Statistical analysis in Small-N Designs: using linear mixed-effects modeling for evaluating intervention effectiveness. Aphasiology 2019;33(1):1-30. [doi: 10.1080/02687038.2018.1454884] [Medline: 33012945]
65. Lugones-Sanchez C, Recio-Rodriguez JI, Agudo-Conde C, et al. Long-term effectiveness of a smartphone app combined with a smart band on weight loss, physical activity, and caloric intake in a population with overweight and obesity (evident 3 study): randomized controlled trial. J Med Internet Res 2022 Feb 1;24(2):e30416. [doi: 10.2196/30416] [Medline: 35103609]
66. Yoshimura E, Tajiri E, Michiwaki R, Matsumoto N, Hatamoto Y, Tanaka S. Long-term effects of the use of a step count-specific smartphone app on physical activity and weight loss: randomized controlled clinical trial. JMIR Mhealth Uhealth 2022 Oct 24;10(10):e35628. [doi: 10.2196/35628] [Medline: 36279159]
67. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. Inquiry 2022;59. [doi: 10.1177/00469580221096272] [Medline: 35471138]
68. Hall KS, Hyde ET, Bassett DR, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. Int J Behav Nutr Phys Act 2020 Jun 20;17(1):78. [doi: 10.1186/s12966-020-00978-9] [Medline: 32563261]
69. Garduno AC, LaCroix AZ, LaMonte MJ, et al. Associations of daily steps and step intensity with incident diabetes in a prospective cohort study of older women: the OPACH Study. Diabetes Care 2022 Feb 1;45(2):339-347. [doi: 10.2337/dc21-1202] [Medline: 35050362]
70. Matz-Costa C, Howard EP, Castaneda-Sceppa C, Diaz-Valdes Iriarte A, Lachman ME. Peer-based strategies to support physical activity interventions for older adults: a typology, conceptual framework, and practice guidelines. Gerontologist 2019 Nov 16;59(6):1007-1016. [doi: 10.1093/geront/gny092] [Medline: 30085074]
71. Mead S, Hilton D, Curtis L. Peer support: a theoretical perspective. Psychiatr Rehabil J 2001;25(2):134-141. [doi: 10.1037/h0095032] [Medline: 11769979]
72. Ikegami S, Takahashi J, Uehara M, et al. Physical performance reflects cognitive function, fall risk, and quality of life in community-dwelling older people. Sci Rep 2019 Aug 22;9(1):12242. [doi: 10.1038/s41598-019-48793-y] [Medline: 31439872]
73. Sanchez M, Vidal JS, Bichon A, et al. Impact of a public open-access community-based physical activity and fall prevention program on physical performance in older adults. Eur J Public Health 2023 Feb 3;33(1):132-138. [doi: 10.1093/eurpub/ckac186] [Medline: 36594685]
74. Buchman AS, Leurgans SE, Wang T, et al. Motor function is the primary driver of the associations of sarcopenia and physical frailty with adverse health outcomes in community-dwelling older adults. PLoS One 2021;16(2):e0245680. [doi: 10.1371/journal.pone.0245680] [Medline: 33529220]
75. Sawada S, Ozaki H, Natsume T, et al. The 30-s chair stand test can be a useful tool for screening sarcopenia in elderly Japanese participants. BMC Musculoskelet Disord 2021 Jul 24;22(1):639. [doi: 10.1186/s12891-021-04524-x] [Medline: 34303339]
76. Pilotto A, Custodero C, Maggi S, Polidori MC, Veronese N, Ferrucci L. A multidimensional approach to frailty in older people. Ageing Res Rev 2020 Jul;60:101047. [doi: 10.1016/j.arr.2020.101047] [Medline: 32171786]
77. Doi T, Tsutsumimoto K, Nakakubo S, et al. Physical performance predictors for incident dementia among Japanese community-dwelling older adults. Phys Ther 2019 Sep 1;99(9):1132-1140. [doi: 10.1093/ptj/pzz077] [Medline: 31162614]
78. Arai H, Ouchi Y, Toba K, et al. Japan as the front-runner of super-aged societies: perspectives from medicine and medical care in Japan. Geriatr Gerontol Int 2015 Jun;15(6):673-687. [doi: 10.1111/ggi.12450] [Medline: 25656311]
79. Tudor-Locke C, Craig CL, Aoyagi Y, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011 Jul 28;8(1):80. [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
80. Nakazono T, Kamide N, Ando M. The reference values for the chair stand test in healthy Japanese older people: determination by meta-analysis. J Phys Ther Sci 2014 Nov;26(11):1729-1731. [doi: 10.1589/jpts.26.1729] [Medline: 25435687]
81. Hoehle H, Zhang X, Venkatesh V. An espoused cultural perspective to understand continued intention to use mobile applications: a four-country study of mobile social media application usability. Eur J Inf Syst 2015 May;24(3):337-359. [doi: 10.1057/ejis.2014.43]
82. Gera R, Chadha P, Ahuja V. Mobile app usage and adoption: a literature review. Int J Electron Bus 2020;15(2):160. [doi: 10.1504/IJEB.2020.106546]
83. Wang C, Qi H. Influencing factors of acceptance and use behavior of mobile health application users: systematic review. Healthcare (Basel) 2021 Mar 22;9(3):357. [doi: 10.3390/healthcare9030357] [Medline: 33809828]

Abbreviations

BMI: body mass index
CS-30: 30-second chair stand test
DPSA: digital peer-supported app
IQR: interquartile range
LPA: light-intensity physical activity
MET: metabolic equivalent
MVPA: moderate to vigorous intensity physical activity
RCT: randomized controlled trial
SB: sedentary behavior

Edited by Q Fan; submitted 26.09.24; peer-reviewed by J Smith-Turchyn, W Liang; revised version received 13.04.25; accepted 14.04.25; published 28.05.25.

Please cite as:
Tabira K, Oguma Y, Yoshihara S, Shibuya M, Nakamura M, Doihara N, Hirata A, Manabe T, Yamashita T
A 12-Month Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Follow-Up Study of a Nonrandomized Controlled Trial
JMIR Aging 2025;8:e66610
URL: https://aging.jmir.org/2025/1/e66610
doi:10.2196/66610

Sexual Response Problems and Their Correlates Among Older Adults From the Sexual Well-Being (SWELL) Study in China: Multicenter Cross-Sectional Study

Bingyu Liang[1,2*], PhD; Chen Xu[3*], PhD; Bingyi Wang[4*], PhD; Xinyi Li[4*], BM; Xin Peng[4*], MD; Ying Wang[5], PhD; Hui Li[6], BM; Yong Lu[7], PhD; Xiaopei Shen[8], MD; Lin Ouyang[9], MD; Guohui Wu[9], MD; Maohe Yu[10], MD; Jiewei Liu[11], BM; Xiaojun Meng[12], PhD; Yong Cai[3*], PhD; Huachun Zou[4,13*], PhD

[1]School of Public Health, Guangxi Key Laboratory of AIDS Prevention and Treatment, Guangxi Medical University, Nanning, China

[10]Department of AIDS/STD Control and Prevention, Tianjin Center for Disease Control and Prevention, Tianjin, China

[11]Baiyun District Center for Disease Control and Prevention, Guanghzou, China

[12]Department of AIDS/STD Control and Prevention, Wuxi Municipal Center for Disease Control and Prevention, Wuxi, China

[13]School of Public Health, Fudan University, 130 Andong Road, Xuhui District, Shanghai, China

[2]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[3]Public Health Research Center, Hongqiao International Institute of Medicine, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

[5]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[6]Department of AIDS/STD Control and Prevention, Shizhong District Center for Disease Control and Prevention, Jinan, China

[7]School of Public Health, the Key Laboratory of Environmental Pollution Monitoring and Disease Control, Guizhou Medical University, Guizhou, China

[8]Department of AIDS/STD Control and Prevention, Affiliation Shanghai Municipal Center for Disease Control and Prevention, Shanghai, China

[9]Department of AIDS/STD Control and Prevention, Chongqing Municipal Center for Disease Control and Prevention, Chongqing, China

[*]these authors contributed equally

Corresponding Author:
Huachun Zou, PhD
School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

Abstract

Background: Sexual response problems among older adults are not an inevitable consequence of aging but rather a response to sexual health. However, there is a lack of recent and multicenter data on this issue in China.

Objective: This study aims to assess the prevalence of sexual response problems and their correlates among older adults.

Methods: A multicenter cross-sectional study on sexual well-being was conducted among individuals aged more than 50 years in China between June 2020 and December 2022. Data on sociodemographics, physical health, psychological health, and sexual response problems were collected through face-to-face interviews. We included sexually active older adults who reported either vaginal, oral, or anal sex in the past 12 months for this study. Sexual response problems included a lack of interest or enjoyment in sex; feeling anxious, having pain, or no excitement during sex; no desire or orgasms; and the lack of lubrication in sex. The stepwise logistic regression models were used to examine the correlates of sexual response problems.

Results: A total of 1317 sexually active older adults (842 men, 475 women) were included. Older women reported a higher prevalence of sexual response problems than older men (52.0% [247/475] vs 43.1% [363/842]). Common factors associated with at least one of the sexual response problems included living in rural areas (men: adjusted odds ratio [aOR]=0.31, 95% CI 0.22 - 0.43; women: aOR=0.29, 95% CI 0.19 - 0.43) and abnormal BMI (aOR=men: 1.52, 95% CI1.11 - 2.07; women: aOR=2.19, 95% CI 1.47 - 3.28). Among older men, sleep quality (aOR=1.87, 95% CI 1.30 - 2.68), emotional connection with sex partners during sexual intercourse (aOR=0.69, 95% CI 0.50 - 0.96), frequently experienced fatigue (aOR=2.47, 95% CI 1.59 - 3.90), anxiety (aOR=4.26, 95% CI 1.12 - 21.27), and seeking professional help for sex life (aOR=1.58, 95% CI 1.14 - 2.21) were associated with sexual response problems. Among older women, sexual response problems were associated with a lack of physical exercise (aOR=1.69, 95% CI 1.13 - 2.54), poor sex-partner relationships (aOR=1.70, 95% CI 1.12 - 2.60), and depressive symptoms (aOR=3.18, 95% CI 1.18 - 10.24).

Conclusions: Sexual response problems are common among older adults. These problems were associated with adverse physical health, mental health, and poor sex-partner relationships. These findings highlight the importance for health care providers to take into account the physical and psychological health of older adults, as well as the quality of their relationships with sexual partners when diagnosing and addressing sexual response problems.

(JMIR Aging 2025;8:e66772) doi:10.2196/66772

KEYWORDS

dysfunction; sexual health; sexual well-being; sexually active; sexual activity; well-being; correlate; sex partner relationship; gerontology; geriatrics; older adults; elder; elderly; older person; aging; China; cross-sectional study

Introduction

Sexual response problems are characterized by diminished or absent sexual interest and disruptions in the physiological or psychosocial patterns associated with the sexual response cycle. These disruptions include a lack of interest, enjoyment, desire, orgasms, and lubrication in sex [1]. Sexual response problems have a profound impact on various aspects of life for older adults, affecting their quality of life, interpersonal relationships, dissatisfaction in marriage [2,3], work efficiency [4], self-esteem [5], and physical and mental health [2,3,6].

Recent studies have shown a high prevalence of sexual problems in both older men and older women. For instance, in Europe, a majority of men (73.7% - 79.8%) and women (23.5% - 50.2%) reported experiencing at least one sexual problem, with the most prevalent sexual problems being orgasmic difficulties and reaching orgasm more quickly than desired and failure to reach orgasm or taking too long to climax [7,8]. An early national survey in urban China indicated that 21% of men and 35% of women aged 20 - 64 years had at least one persistent sexual dysfunction, with lack of sex interest, erectile difficulties (men), and inability to reach orgasm being the most common problems [9]. There are substantial variations in the prevalence of sexual response problems across different global regions, with noticeable differences between men and women. As extensively demonstrated by a significant number of epidemiological studies worldwide, there are substantial differences in the prevalence of sexual response problems among men and women. A few researchers have suggested that the prevalence of sexual response problems was higher in men compared to women [10,11]. Conversely, more studies have reported a higher prevalence among females [12-15]. It is necessary to elucidate the specific causes of these differences and inform gender-specific prevention and intervention strategies.

Sociodemographic, psychological, sex partner relationships, chronic disease, and physical health were found to be important determinants of sexual response problems among older adults. The prevalence of sexual problems tends to increase with age [16], and there are notable differences between men and women in terms of prevalence and types of sexual response problems [17]. Psychosocial factors such as anxiety, depression, stress, and the quality of marital relationships play significant roles in sexual response problems [16,18,19]. Intimate partner violence is associated with a higher likelihood of experiencing sexual problems [20]. In addition, chronic diseases like cancer [17], diabetes [21], and coronary heart disease [22] are linked to an increased risk of sexual dysfunction. Moreover, sexual response problems in one partner may influence the sexual function of the other partner [23].

Despite extensive studies on sexual problems, most of the existing literature was from developed countries or early studies, and there is a lack of nationally representative, large sample and recent data on the prevalence and correlates of sexual response problems among older adults in China. Given China's rapidly aging population [24,25], the sexual health of older adults is a growing concern. A comprehensive understanding of older adults' sexual response problems may enhance sex education, research, policy, and clinical care for this growing population. This multicentre cross-sectional study, using data from the Sexual Well-being (SWELL) study in China, aims to fill the research gap by examining the prevalence of sexual response problems and their correlates among older adults. These epidemiological data are essential for andrologists, gynecologists, urologists, venereologists, and other health care providers involved in treating and caring for older adults. They may help them counsel their patients on the potential adverse effects of different treatment modalities. Moreover, the findings are expected to contribute valuable insights for developing targeted interventions to enhance sexual relationships, improve quality of life, and address the sexual health challenges faced by aging populations in China.

Methods

Study Participants and Procedures

Our study was based on the SWELL study, a multicenter cross-sectional survey conducted between June 2020 and December 2022. The survey spanned four different regions in China, including Shanghai (Eastern China), Jinan (Eastern China), Chongqing (Western China), Guangzhou (Southern China), and Tianjin (Northern China). Participants were recruited using a multistage sampling design, and more detailed sampling procedures are provided in our previous protocol [26]. We collected data on demographic characteristics, physical health characteristics, mental health characteristics, sex partner relationship characteristics, and sexual behavior characteristics through face-to-face interviews. All participants provide formal informed consent to participate in the study.

Participants were enlisted from subdistricts within each chosen city. Eligibility criteria for participants in this study included: (1) aged 50 years and older; (2) only heterosexual orientation; (3) having engaged in heterosexual activities (including oral or vaginal intercourse) in the preceding year; (4) being able to comprehend the survey instrument of the SWELL Study.

Ethical Consideration

The SWELL study was approved by the School of Public Health (Shenzhen), Sun Yat-sen University Research Ethics Committee (approval number SYSU-PHS [2019] 006) and was performed following the Helsinki Declaration. Written consent was obtained from all participants, who were informed of their right to withdraw from the study at any time. Participant information and responses remained confidential, with anonymized data stored in password-protected folders accessible exclusively to the research team and supervisors.

Study Variables

Sexual Response Problems (Outcomes Variables)

Respondents were asked to report if they had experienced any of the following sexual response problems for three months in the preceding year: (1) lacked desire for sex, (2) lacked enjoyment in sex, (3) anxiety during sex, (4) discomfort or pain in sex, (5) no sexual arousal or excitement during sex, (6) lack of or delayed orgasm despite arousal, and (7) reaching orgasm faster than you would like, (8) lubrication difficulties (women only) or erectile function difficulties (men only). These items captured the major sexual response problem domains in the classification of sexual dysfunction in the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition [27]. The reliability of sexual response problems in this study was 0.75.

Sociodemographic Characteristics

Age, gender (men and women), living area (rural or urban), monthly income (Chinese Yuan [CNY]), and years of education (<6 years: primary and lower; 7 - 12 years: senior or high school; >12 years: college and above) were included in demographic characteristics. Age was categorized into 3 age groups (50 - 59 years, 60 - 69 years, and older than 70 years). Monthly income (CNY) was categorized into 2 groups (≥5000 CNY [US $700] and <5000 CNY [US $700]).

Lifestyle Characteristics

Physical Exercise

Physical exercise was assessed with 5 items that inquired about the frequency of participation (6 times a week, 3 - 5 times a week, 1 - 2 times a week, no more than once a week, hardly ever, or never). The participants who reported engaging in physical exercise more than 1 - 2 times a week were categorized as often engaging in physical exercise. The remaining responses were categorized as not often exercised [28].

Seeking Professional Help for Sex Life

Participants were asked whether they had sought help or advice about their sex life from a range of sources in the past year. These sources included 4 informal sources (family member or friend, information and support sites on the internet) and 6 professional sources (general practitioner or family doctor, sexual health clinic, genitourinary clinic, sexually transmitted infection clinic, or relationship counselor); more than one answer was allowed. If a participant responds that they have previously sought help for sex life from 6 professional sources, this is defined as “seeking professional help for sex life.”

Physical Health Characteristics

Frequently Experiencing Fatigue

As for frequently experiencing fatigue, the participants were asked about their fatigue levels using a question: “Do you frequently experience fatigue?” with response options of “Yes” or “No.”

Chronic Disease

Chronic disease is defined as one or more diseases that last for 3 months or more including cardiovascular diseases (including myocardial infarction, coronary heart disease, angina, other forms of heart disease, and hypertension), arthritis, diabetes, or hyperglycemia, cerebral hemorrhage or stroke, chronic lung disease (excluding asthma), Parkinson disease, epilepsy, hyperlipidemia, gout or hyperuric acid, chronic gastroenteritis and chronic pain. A separate section was directed at women participants, whether they have a history of one of the following, including bladder surgery, genital or gynecologic surgery, cesarean section, abdominal surgery, and hip pelvic fractures or hip replacement.

Body Mass Index and Sleep Quality

In the SWELL Study, BMI is considered abnormal when lower than 18.5 or higher than 25.

Sleep quality was evaluated using the validated single-item sleep quality scale (SQS), which ranges from 1 to 10 and has been proven to be divided into 2 categories for analysis: those indicating poor sleep quality (scores 1 - 6) and those indicating good sleep quality (scores 7 - 10) [29]. The SQS had an acceptable internal consistency (Cronbach α=0.85).

Mental Health Characteristics

Depressive Symptoms

Depressive symptoms were measured by the 9-item Patient Health Questionnaire (PHQ-9), which has been validated and proven [29,30]. The scale’s total score ranges from 0 to 27, with scores≥10 representing clinically significant depressive symptoms. This study defined a score greater than or equal to 10 as depressive symptoms.

Anxiety Symptoms

The anxiety symptoms were measured on the generalized anxiety disorder-7 (GAD-7) scale with 7-item. Mild or normal anxiety was defined as a GAD-7 score <10, while moderate-to-vigorous anxiety was defined as a GAD-7 score ≥10 [29]. In this study, a score greater or equal to 10 was defined as with anxiety symptoms.

Sexual Relationship Characteristics

Emotional Connection With a Sexual Partner During Sex

The participants were asked, “How often would you say you feel emotionally close to your partner when you have sex together? (options: always, most of the time, sometimes, not very often, hardly ever).” The participants who selected the options of “always” and “most of the time” were defined as having a “good emotional connection with a sexual partner

during sex;" otherwise, they were defined as having a "poor emotional connection with a sexual partner during sex."

Relationship With Sex Partners

Regarding relationships with sex partners, the participants were asked, "How do you evaluate the relationship with your recent sex partner? Please assign a rating to the quality of your partnership with them." The rating scale ranges from 1 to 7, with 1 indicating "very good" and 7 representing "very poor." We categorized the responses into 2 categories for analysis: good relationship with a sex partner (scores 1 - 4) and poor relationship with a sex partner (scores 5 - 7).

Sexual Satisfaction

Sexual satisfaction was measured on a 5-point Likert scale, with responses ranging from 1 to 5 (strongly agree, agree, medium, disagree, and strongly disagree). In this study, we reclassified scores of 1 - 3 as sexual satisfaction and scores of 4 - 5 as sexual dissatisfaction [30].

Statistical Analyses

Descriptive analyses were conducted to characterize the study sample, including presenting percentages, means, and SD. The χ^2 test was used to compare the proportions of characteristics between the sex groups.

For the multivariable logistic regression analysis, collinearity diagnostics were initially performed for all potential variables (Multimedia Appendices 1 and 2). Subsequently, multivariate logistic regression analysis was carried out for noncollinear variables. The multivariable logistic regression models selected variables using a stepwise method based on the Akaike Information Criterion (AIC). The stepwise regression method combines forward selection and backward elimination approaches, adding and removing predictors in the model-building process. This approach effectively minimizes the inclusion of covariates, thereby enhancing the robustness of the analysis. Finally, the model with the minimum AIC was adopted (men: 1005.929; women: 582.316). Adjusted odds ratios (aORs) and their corresponding 95% CIs were estimated.

All statistical analyses were performed using R software version 4.2.3 (R Project). The Stats package (version 4.2.2) was used to build the stepwise multivariable logistic regression models. In addition, the figures were generated using the ggplot2 package (version 3.4.3) and forestmodel package (version 0.6.2) from CRAN.

Results

Demographic and Health Characteristics of the Participants

As shown in Table 1, 1317 older adults were included in this analysis. The average age was 64 years (SD 8.4 years, ranging from 50 to 86 years). Over half of the participants resided in rural areas (men: 53.1% [447/842], women: 50.9% [242/475]), and the majority reported 7 - 12 years of education (junior or senior high school; men: 70.0% [589/842], women: 59.6% [283/475]). In addition, a significant proportion of participants reported infrequent engagement in physical exercise (men: 54.8% [461/842], women: 48.2% [229/475]). Regarding physical health, more than half of the participants did not frequently experience fatigue (men: 85% [716/842], women: 74.9% [356/475]) and did not have chronic diseases (men: 55.2% [465/842], women: 62.9% [299/475]). Regarding sexual relationship characteristics, the majority of male participants reported sexual satisfaction (479/842, 56.9%) and a good relationship with their sex partner (615/842, 73.0%). In comparison, women reported slightly lower rates of sexual satisfaction (228/475, 48.0%) and a good relationship with their sex partner (306/475, 64.4%). Furthermore, less than 5.1% (24/475) of both men and women reported symptoms of anxiety and depression.

Table . Demographic, lifestyle, health, and sexual relationship characteristics among older adults aged more than 50 years in China (stratified by sex).

Characteristics	Men				Women			
				P values				*P* values
At least one of the sexual response problems, n (%)	All	Yes	No		All	Yes	No	
	842 (100.0)	363 (43.1)	479 (56.9)	N/A[a]	475 (100.0)	247 (52.0)	228 (48.0)	N/A
Demographic characteristics								
Living area				<.001				<.001
Rural	447 (53.1)	132 (29.5)	315 (70.5)		242 (50.9)	89 (36.8)	153 (63.2)	
Urban	395 (46.9)	231 (58.5)	164 (41.5)		233 (49.1)	158 (67.8)	75 (32.2)	
Age (years)				.04				.13
50 - 59	440 (52.3)	175 (39.8)	265 (60.2)		272 (57.3)	137 (50.4)	135 (49.6)	
60 - 69	300 (35.6)	134 (44.7)	166 (55.3)		156 (32.8)	79 (50.6)	77 (49.4)	
70+	102 (12.1)	54 (52.9)	48 (47.1)		47 (9.9)	31 (66)	16 (34)	
Monthly income (RMB)				.37				.72
≥5000	229 (27.2)	105 (45.9)	124 (54.1)		398 (83.8)	205 (51.5)	193 (48.5)	
<5000	613 (72.8)	258 (42.1)	355 (57.9)		77 (16.2)	42 (54.5)	35 (45.5)	
Education level				.009				.02
≤6	160 (19.0)	75 (46.9)	85 (53.1)		121 (25.5)	51 (42.1)	70 (57.9)	
7 - 12	589 (70.0)	236 (40.1)	353 (59.9)		283 (59.6)	152 (53.7)	131 (46.3)	
>12	93 (11.0)	52 (55.9)	41 (44.1)		71 (14.9)	44 (62)	27 (38)	
Lifestyle characteristics								
Physical exercise				.81				.001
Often	381 (45.2)	162 (42.5)	219 (57.5)		246 (51.8)	109 (44.3)	137 (55.7)	.
Not Often	461 (54.8)	201 (43.6)	260 (56.4)		229 (48.2)	138 (60.3)	91 (39.7)	
Seeking professional help for sex life				.001				.96
Yes	266 (31.6)	138 (51.9)	128 (48.1)		138 (29.1)	71 (51.4)	67 (48.6)	
No	576 (68.4)	225 (39.1)	351 (60.9)		337 (70.9)	176 (52.2)	161 (47.8)	
Physical health characteristics								
Frequently experienced fatigue				<.001				.07
Often	126 (15.0)	86 (68.3)	40 (31.7)		119 (25.1)	71 (59.7)	48 (40.3)	.07
Not often	716 (85.0)	277 (38.7)	439 (61.3)		356 (74.9)	176 (49.4)	180 (50.6)	
BMI				.001				<.001
Normal	434 (51.5)	162 (37.3)	272 (62.7)		230 (48.4)	95 (41.3)	135 (58.7)	
Abnormal	408 (48.5)	201 (49.3)	207 (50.7)		245 (51.6)	152 (62)	93 (38)	
Sleep quality				<.001				.35

Characteristics	Men				Women			
				P values				*P* values
Good	630 (74.8)	232 (36.8)	398 (63.2)		314 (66.1)	158 (50.3)	156 (49.7)	
Poor	212 (25.2)	131 (61.8)	81 (38.2)		161 (33.9)	89 (55.3)	72 (44.7)	
Chronic disease				.10				.06
Yes	377 (44.8)	163(43.2)	214 (56.8)		176 (37.1)	81 (46)	95 (54)	
No	465 (55.2)	200 (43)	265 (57)		299 (62.9)	166 (55.5)	133 (44.5)	
Mental health characteristics								
Anxiety symptoms				.002				.34
Yes	17 (2.0)	14 (82.4)	3 (17.6)		20 (4.2)	13 (65)	7 (35)	
No	825 (98.0)	349 (42.3)	476 (57.7)		455 (95.8)	234 (51.4)	221 (48.6)	
Depressive symptoms				<.001				.012
Yes	37 (4.4)	30 (81.1)	7 (18.9)		24 (5.1)	19 (79.2)	5 (20.8)	
No	805 (95.6)	333 (41.4)	472 (58.6)		451 (94.9)	228 (50.6)	223 (49.4)	
Sexual relationship characteristics								
Emotional connection with a sexual partner during sex				<.001				.004
Yes	573 (68.1)	223 (38.9)	350 (61.1)		294 (61.9)	144 (49)	150 (51)	
No	269 (31.9)	140 (52.0)	129 (48.0)		181 (38.1)	103 (56.9)	78 (43.1)	
Relationship with a sex partner				.002				.005
Good	615 (73.0)	245 (39.8)	370 (60.2)		306 (64.4)	144 (47.1)	162 (52.9)	
Poor	227 (27.0)	118 (52)	109 (48)		169 (35.6)	103 (60.9)	66 (39.1)	
Sexual satisfaction[b]				<.001				<.001
Yes	479 (56.9)	0 (0)	479 (100)		228 (48.0)	0 (0)	228 (100)	
No	363 (43.1)	363 (100)	0 (0)		247 (52.0)	247 (100)	0 (0)	

[a]N/A: not applicable

[b]Sexual response problems include lacked interest in having sex, lacked enjoyment in sex, feeling anxiety during sex, feeling physical pain as a result of sex, feeling no excitement or arousal during sex, difficulty in reaching climax, reaching climax more quickly than desired, trouble getting or keeping an erection(men) or uncomfortable dry vagina(women).

Prevalence of Sexual Response Problems

The prevalence of at least one sexual response problem (including or excluding lack of interest in sex) is shown in Figure 1 and Table 2. In total, 610 out of 1317 participants had sexual response problems, with an overall prevalence of sexual response problems of 46.3% (610/1317). There was a significant difference in the prevalence of the reported at least one sexual response problem between women and men, with being significantly higher in women than in men (52.0% [247/475] vs 43.1% [363/842], χ^2_1=9.6, *P*=.002). The prevalence of at least one sexual response problem increased with age among men (The Cochran-Armitage Trend Test, Z=−2.476, *P* for trend=.01). However, this trend was only observed in the age group between the older than 70 years age group and the other 2 age groups among men. There was no significant difference in the prevalence of sexual response problems between the 60 - 69 years age group and the 50 - 59 age group among women.

Figure 1. Prevalence of sexual response problems among older adults aged 50+ years**.** (**A**) At least one of sexual response problems among men and women; (**B**) At least one of sexual response problems excluding lack of interest in sex among men and women.

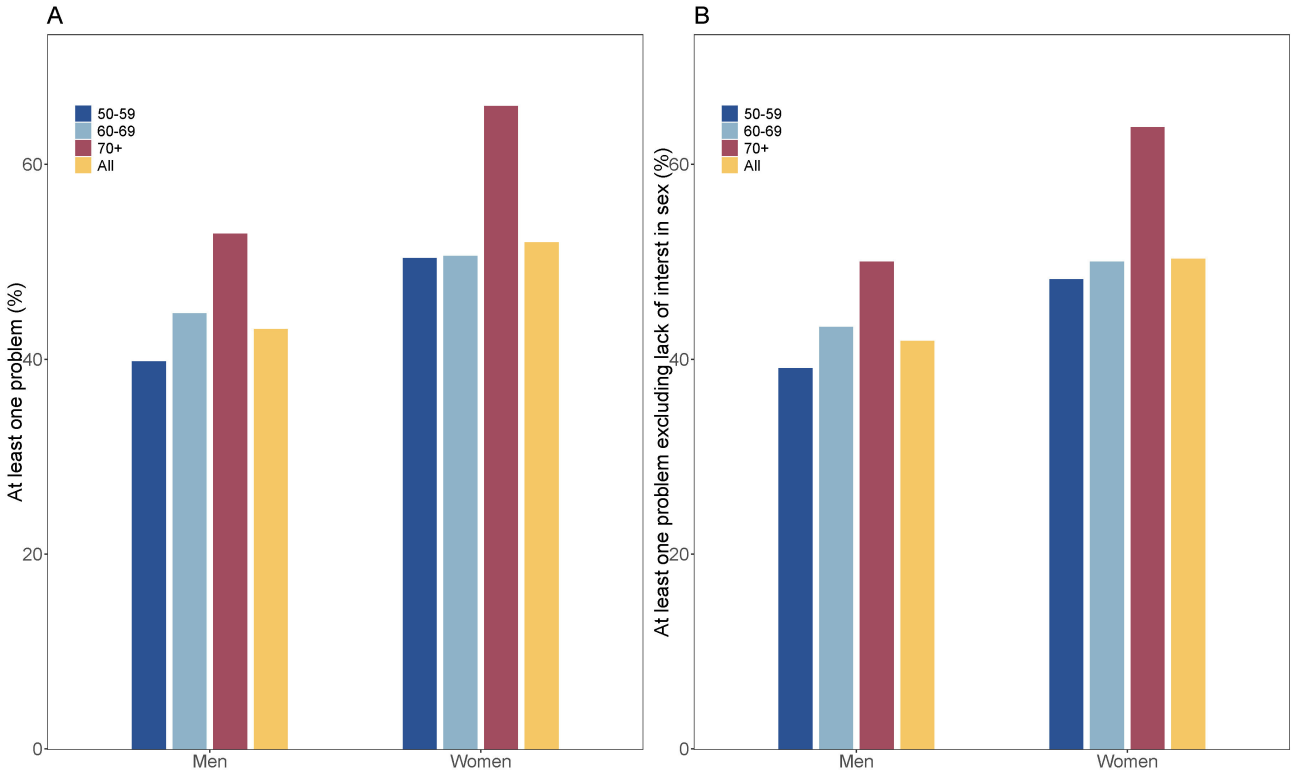


Table . The prevalence of sex response problems among men and women by age group.

Age group	Men	Women	Chi-square (df)	*P* values
At least one sex response problem				
50 - 59	175 (39.8)	137 (50.4)	7.24 (1)	.007
60 - 69	134 (44.7)	79 (50.6)	1.24 (1)	.27
Older than 70 years	54 (52.9)	31 (66.0)	1.73 (1)	.19
All age groups	363 (43.1)	247 (52.0)	9.65 (1)	.002
At least one sex response problem excluding lack of interest in sex				
50 - 59 years	172 (39.1)	131 (48.2)	5.29 (1)	.02
60 - 69 years	130 (43.3)	78 (50.0)	1.58 (1)	.21
Older than 70 years	51 (50)	30 (63.8)	1.95 (1)	.16
All age groups	353 (41.9)	239 (50.3)	8.31 (1)	.004

Correlates of Sexual Response Problems

The results of multivariable logistic regression analysis stratified by sex, which was presented in Figure 2, revealed several significant associations with reporting at least one sexual response problem.

Figure 2. Correlates of sexual response problems among older adults aged 50+ years**.** (**A**) Correlates of sexual response problems among men aged 50+ years; the multivariable logistic regression analysis was adjusted by living area, age, educational level, feeling of fatigue, body mass index, sleeping quality, seeking professional help for sex life, emotional connection with sexual partner during sex, anxiety, depressive symptoms. (**B**) Correlates of sexual response problems among women aged 50+ years; the logistic regression model was adjusted by age, living area, physical exercise, body-mass index, sleeping quality, sexual relationship with partner, emotional connection with sexual partner during sex, depressive symptoms. SRP: Sexual Response Problem; aOR: adjusted odds ratio; CI: confidence interval.

A

Variables	Level	N *(%)*	Prevalence of *SRP (%)*	*Adjusted odds ratio*	*aOR (95%CI)*	*p*
Living area	Urban	395 (46.91)	231 (58.48)		Reference	
	Rural	447 (53.09)	132 (29.53)		**0.31 (0.22, 0.43)**	**<0.001**
Age	50-59	440 (52.26)	175 (39.77)		Reference	
	60-69	300 (35.63)	134 (44.67)		1.28 (0.91, 1.79)	0.155
	70+	102 (12.11)	54 (52.94)		**1.72 (1.05, 2.81)**	**0.030**
Education level	Primary school and lower	93 (11.05)	52 (55.91)		Reference	
	Senior or high school	160 (19.00)	75 (46.88)		1.44 (0.78, 2.66)	0.241
	College and above	589 (69.95)	236 (40.07)		0.78 (0.47, 1.29)	0.336
Feeling of fatigue	No	716 (85.04)	277 (38.69)		Reference	
	Yes	126 (14.96)	86 (68.25)		**2.47 (1.59, 3.90)**	**<0.001**
Body-mass index	Normal	434 (51.54)	162 (37.33)		Reference	
	Abnormal	408 (48.46)	201 (49.26)		**1.52 (1.11, 2.07)**	**0.008**
Sleep quality	Good	630 (74.82)	232 (36.83)		Reference	
	Bad	212 (25.18)	131 (61.79)		**1.87 (1.30, 2.68)**	**<0.001**
Seeking professional help for sex life	No	576 (68.41)	225 (39.06)		Reference	
	Yes	266 (31.59)	138 (51.88)		**1.58 (1.14, 2.21)**	**0.007**
Emotional connection with sex partner during sex	No	269 (31.95)	140 (52.04)		Reference	
	Yes	573 (68.05)	223 (38.92)		**0.69 (0.50, 0.96)**	**0.029**
Anxiety (moderate to severe anxiety)	No	825 (97.98)	349 (42.30)		Reference	
	Yes	17 (2.02)	14 (82.35)		**4.26 (1.12, 21.27)**	**0.048**
Depressive symptoms	No	805 (95.61)	333 (41.37)		Reference	
	Yes	37 (4.39)	30 (81.08)		2.32 (0.94, 6.44)	0.083

B

Variables	Level	N *(%)*	Prevalence of *SRP (%)*	*Adjusted odds ratio*	*aOR (95%CI)*	*p*
Living area	Urban	233 (49.05)	158 (67.81)		Reference	
	Rural	242 (50.95)	89 (36.78)		**0.29 (0.19, 0.43)**	**<0.001**
Age	50-59	272 (57.26)	137 (50.37)		Reference	
	60-69	156 (32.84)	79 (50.64)		0.90 (0.58, 1.39)	0.62
	70+	47 (9.89)	31 (65.96)		1.93 (0.96, 3.99)	0.07
Physical exercise	Yes	246 (51.79)	109 (44.31)		Reference	
	No	229 (48.21)	138 (60.26)		**1.69 (1.13, 2.54)**	**0.01**
Body-mass index	Normal	230 (48.42)	95 (41.30)		Reference	
	Abnormal	245 (51.58)	152 (62.04)		**2.19 (1.47, 3.28)**	**<0.001**
Sleep quality	Good	314 (66.11)	158 (50.32)		Reference	
	Bad	161 (33.89)	89 (55.28)		1.42 (0.93, 2.19)	0.11
Relationship with sex partner	Good	306 (64.42)	144 (47.06)		Reference	
	Bad	169 (35.58)	103 (60.95)		**1.70 (1.12, 2.60)**	**0.01**
Emotional connection with sex partner during sex	No	181 (38.11)	103 (56.91)		Reference	
	Yes	294 (61.89)	144 (48.98)		0.88 (0.57, 1.34)	0.54
Depressive symptoms	No	451 (94.95)	228 (50.55)		Reference	
	Yes	24 (5.05)	19 (79.17)		**3.18 (1.18, 10.24)**	**0.03**

For older men, residing in rural areas (aOR 0.31, 95% CI 0.22 - 0.43) and maintaining an emotional connection with a sex partner during sexual intercourse (aOR 0.69, 95% CI 0.50 - 0.96) were associated with a reduced likelihood of reporting sexual response problems. In contrast, being 70 years or older (aOR 1.72, 95% CI 1.05 - 2.81), frequently experiencing fatigue (aOR 2.47, 95% CI 1.59 - 3.90), and poor sleeping quality (aOR 1.87, 95% CI 1.30 - 2.68) were associated with higher odds of reporting sexual response problems. In addition, moderate to severe anxiety symptoms (aOR 4.26, 95% CI 1.12 - 21.27), abnormal BMI (aOR 1.52, 95% CI 1.11 - 2.07), and seeking professional help for sex issues (aOR 1.58, 95% CI 1.14 - 2.21) were positively associated with reporting sexual response problems among older men.

Older women residing in rural areas (aOR 0.29, 95% CI 0.19 - 0.43) were less likely to report sexual response problems. Conversely, factors such as infrequent engagement in physical exercise (aOR 1.69, 95% CI 1.13 - 2.54), depressive symptoms (aOR 3.18, 95% CI 1.18 - 10.24), having an abnormal BMI (aOR 2.19, 95% CI 1.47 - 3.28), and being in a poor sex partner relationship (aOR 1.70, 95% CI 1.12 - 2.60) were associated with higher odds of reporting sexual response problems.

Discussion

Principal Results

Our study demonstrated a high prevalence of sexual response problems among older adults, with variations noted between men and women. In older men, sexual response problems correlate with advancing age. Our findings also linked sexual response problems with adverse physical health outcomes such as frequent experiencing fatigue, poor sleep quality, and abnormal BMI. In addition, we observed strong associations between sexual response problems and mental health issues, including anxiety and depressive symptoms. Moreover, poor sex partner relationships, sexual dissatisfaction, and lack of emotional connection during sex were also associated with sexual response problems. These findings significantly contribute to the existing literature, underscoring the importance of addressing sexual response problems within the domain of sexual health and enhancing our understanding of these issues among older adults.

Comparison With Previous Work

In this study, women reported a higher prevalence of sexual response problems compared to men, consistent with findings from other countries [12-14]. This discrepancy may be due to several factors, including physiological alterations, psychological elements, and sociocultural dynamics. With advancing age, women experience a significant decline in estrogen levels, especially after menopause [31]. Furthermore, older women are subject to various age-related psychological changes, such as concerns about body image, fears about aging, and self-consciousness about their sexual lives, which may lead to decreased sexual desire or arousal issues. In Chinese culture, the sexual lives of older adults are frequently neglected or deemed inappropriate, which may influence the perceptions and expectations of sex among older women, making them feel ashamed or uncomfortable in their sexual lives. This sex difference highlights the need for targeted interventions for women's sexual health. These interventions should address the multifaceted nature of sexual health in older women, combining physiological, psychological, and sociocultural interventions.

In this study, adverse physical health, such as abnormal BMI, frequent fatigue, and poor sleep quality, played a crucial role in sexual response problems. These findings were identified by evidence from high-income countries, which linked poor physical health to sexual response problems [32-34]. Obesity, characterized by an abnormal BMI, detrimentally impacts the reproductive system and sexual function [35-37]. Both men and women affected by obesity face a heightened risk of fertility challenges and sexual dysfunction [38]. Weight loss can reduce fatty tissue and diminish aromatase activity, leading to a relative increase in testosterone levels [39], which enhances sexual function in men. Besides, weight loss tends to improve sexual functioning for women and men [40]. Obesity significantly affects the hypothalamic-pituitary-gonadal axis in men, leading to diminished libido and erectile dysfunction [41]. The previous review highlighted that excess body weight negatively affected hormones contributing to sexual behavior, noting that adipose tissue facilitates the conversion of androgens to estrogens, further impacting sexual function [42]. Given the inverse association between body mass and sexual response problems, it is recommended that clinicians, both in general practice and in weight loss programs, should more fully address sexual response problems.

Poor sleep quality was associated with sexual response problems in previous studies [43,44], which was also demonstrated by this study. Poor sleep is closely linked to sexual dysfunction due to several physiological and psychological factors. Inadequate sleep can disrupt hormone production, notably reducing testosterone levels, which are essential for sexual desire and performance. Studies have shown that sleep deprivation can decrease testosterone production, leading to impaired sexual activity [45]. In addition, sleep disorders often contribute to stress, anxiety, and depression, all of which can diminish libido and sexual satisfaction [46]. Furthermore, chronic stress and anxiety can lead to sleep problems, which in turn may cause erectile dysfunction [47]. Sleep is fundamental to health, and its bidirectional relationship with sexual response difficulties necessitates that clinicians conduct thorough assessments to identify underlying causes of poor sleep, encourage patient-partner communication to alleviate psychological burdens, promote regular physical activity to enhance sleep quality, and recommend professional sleep therapy to improve physiological function when necessary.

Our study showed that lack of physical exercise was correlated with the occurrence of sex response problems among older women. Regular physical activity is a healthy practice that can mitigate the risk of sexual response problems [36,48]. These associations between sexual response problems and physical health underscore the importance of prioritizing sexual function within sexual health. Physical exercise and high-quality sleep are recommended, along with other lifestyle guidance, to improve sexual functioning in both men and women and to improve health across a range of domains.

Psychosocial factors, including symptoms of anxiety and depression, exhibited the strongest association with sexual response problems, as evidenced by findings from the Global Study of Sexual Attitudes and Behaviours (GSSAB) [49]. Other studies also linked the associations between mental health and sexual response problems, which highlighted that men with anxiety disorders are at a higher risk of developing erectile dysfunction [50,51]. Moreover, the physiological responses triggered by heightened anxiety levels may contribute to disruptions in sexual function [52]. This finding will capture the attention of practitioners exploring the causes and treatment of sexual problems in patients. In clinical practice, there should be heightened efforts to address sexual problems within integrated services of mental health and sexual health counseling, as well as in primary and secondary care.

Our findings underscored the importance of sex-partner relationships in the context of sexual response problems. Specifically, men who maintained an emotional connection with their sex partners during sex were less likely to report sex response problems, while women experiencing poor relationships with sex partners were more likely to encounter sexual response problems in older women. This underscores the role of brief emotional interactions during sexual activity in men's adaptation to sexual response problems. In contrast, long-term sexual relationships appear to be more influential in women's adjustment to sexual response problems. Existing studies have consistently underscored the significance of partner relationships for sexual response problems in women [53], emphasizing the impact of daily intimacies in relationships [9]. In contrast, emotional intimacy is presumed to play a significant role in maintaining sexual desire and partnered sexual activity for men [54,55]. In addition, emotional connection with a partner may foster open communication, trust, and mutual understanding, all essential to a satisfying sexual relationship. The variation in sexual response problems concerning sex partner relationships among older men and women underscores the importance of physicians considering sex differences when diagnosing and treating patients with sexual response problems. In future studies, the sex partner should be included and involved in the evaluation and management to achieve a better intimate relationship in an established couple and avoid sexual response problems.

Limitations

Our study has several limitations that should be considered. First, as a cross-sectional study, it cannot establish temporal order and causal direction. Second, the reliance on self-reporting in the questionnaires, especially for sensitive issues, introduced the possibility of recall biases. Third, the sexual response problems were reported by the participants experiencing them, and the differences in the sensitivity and understanding of the same sexual problem may exist among participants, potentially influenced by factors such as education, age, or living area. Finally, some analyses were based on small cell sizes, particularly for variables like anxiety and depressive symptoms, which may result in unstable estimates.

Conclusion

This study showed a substantial prevalence of sexual response problems among older adults, with women experiencing these problems more often than men. The study identified adverse physical health, poor mental health, and poor relationships with sex partners as factors contributing to increased sexual response problems among older adults. To address these concerns, health care professionals can implement interventions for older adults experiencing sexual response problems, such as enhancing physical health, supporting mental health, improving intimate relationships, and providing educational and cognitive-behavioral interventions. These insights drawn from the latest and representative SWELL study data enhance our understanding of sexual response problems among older adults and have the potential to promote overall health and well-being among the aging population in China.

Acknowledgments

We thank the investigation staff, all participants, and Ms Bowen Liang for her outstanding contribution to this study. This study was supported by the Natural Science Foundation of China International/Regional Research Collaboration Project (grant number 72061137001), and the Natural Science Foundation of China Excellent Young Scientists Fund (grant number 82022064). All funding parties had no role in the study design or data explanation.

Data Availability

The data collected in this study will not be publicly available. However, upon reasonable request, the corresponding author can be contacted for de-identified data.

Authors' Contributions

HZ conceived and designed the study in consultation with BL, CX, BW, XL, and XP. BW, XP, YW, HL, YL, XS, LO, GW, MY, JL, and XM contributed to data collection. BL, CX, BW, XM, WT, and JT contributed to data analysis and presentation. BL, CX, BW, XP, and XL drafted the manuscript, with all authors critically reviewing the paper. All authors approved the final report. BL, YC, and HZ contributed equally as co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Collinearity diagnostics of all potential variables for the multivariable logistic regression analysis (Men). Note that no value was more than 0.3, suggesting no collinearity between variables.
[PNG File, 345 KB - aging_v8i1e66772_app1.png]

Multimedia Appendix 2
Collinearity diagnostics of all potential variables for the multivariable logistic regression analysis (Women). Note that no value was more than 0.3, suggesting no collinearity between variables.
[PNG File, 341 KB - aging_v8i1e66772_app2.png]

References

1. Wang B, Peng X, Fu L, et al. Sexual function and correlates among adults aged 50+ years in China: findings from the sexual well-being (SWELL) study. J Am Geriatr Soc 2024 Apr;72(4):1122-1135. [doi: 10.1111/jgs.18789] [Medline: 38319006]
2. Ceyhan O, Ozen B, Simsek N, Dogan A. Sexuality and marital adjustment in women with hypertension in Turkey: how culture affects sex. J Hum Hypertens 2019 May;33(5):378-384. [doi: 10.1038/s41371-019-0181-3] [Medline: 30804462]
3. Galati MCR, Hollist CS, do Egito JHT, et al. Sexual dysfunction, depression, and marital dissatisfaction among Brazilian couples. J Sex Med 2023 Feb 27;20(3):260-268. [doi: 10.1093/jsxmed/qdac004] [Medline: 36763940]
4. Elterman DS, Bhattacharyya SK, Mafilios M, Woodward E, Nitschelm K, Burnett AL. The quality of life and economic burden of erectile dysfunction. Res Rep Urol 2021;13:79-86. [doi: 10.2147/RRU.S283097] [Medline: 33634039]

5. Lew-Starowicz Z, Czajkowska K. Prevalence of sexual dysfunctions and associated risk factors in Poland. Arch Med Sci 2022;18(4):1031-1040. [doi: 10.5114/aoms.2019.86794] [Medline: 35832714]
6. Atlantis E, Sullivan T. Bidirectional association between depression and sexual dysfunction: a systematic review and meta-analysis. J Sex Med 2012 Jun 1;9(6):1497-1507. [doi: 10.1111/j.1743-6109.2012.02709.x]
7. Hald GM, Graham C, Štulhofer A, Carvalheira A, Janssen E, Træen B. Prevalence of sexual problems and associated distress in aging men across 4 European countries. J Sex Med 2019 Aug;16(8):1212-1225. [doi: 10.1016/j.jsxm.2019.04.017] [Medline: 31255584]
8. Graham CA, Štulhofer A, Lange T, et al. Prevalence and predictors of sexual difficulties and associated distress among partnered, sexually active older women in Norway, Denmark, Belgium, and Portugal. Arch Sex Behav 2020 Nov;49(8):2951-2961. [doi: 10.1007/s10508-020-01742-7] [Medline: 32533517]
9. Parish WL, Laumann EO, Pan S, Hao Y. Sexual dysfunctions in urban china: a population-based national survey of men and women. J Sex Med 2007 Nov;4(6):1559-1574. [doi: 10.1111/j.1743-6109.2007.00596.x] [Medline: 17888072]
10. Yu Z, Niu J, Wang C. The prevalence and risk factors of sexual dysfunction in the elderly in Southern China. Int J Gen Med 2024;17:2355-2360. [doi: 10.2147/IJGM.S462124] [Medline: 38803552]
11. Spatz ES, Canavan ME, Desai MM, Krumholz HM, Lindau ST. Sexual activity and function among middle-aged and older men and women with hypertension. J Hypertens 2013 Jun;31(6):1096-1105. [doi: 10.1097/HJH.0b013e32835fdefa] [Medline: 23640604]
12. Mitchell KR, Mercer CH, Ploubidis GB, et al. Sexual function in Britain: findings from the third national survey of sexual attitudes and lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1817-1829. [doi: 10.1016/S0140-6736(13)62366-1] [Medline: 24286787]
13. Andresen JB, Graugaard C, Andersson M, Bahnsen MK, Frisch M. Sexual inactivity and dysfunction in Denmark: a project SEXUS study. Arch Sex Behav 2022 Nov;51(8):3669-3688. [doi: 10.1007/s10508-022-02348-x] [Medline: 36127559]
14. Quinta Gomes AL, Nobre PJ. Prevalence of sexual problems in Portugal: results of a population-based study using a stratified sample of men aged 18 to 70 years. J Sex Res 2014;51(1):13-21. [doi: 10.1080/00224499.2012.744953] [Medline: 23573897]
15. Ramírez-Santos J, Cristóbal-Cañadas D, Parron-Carreño T, Lozano-Paniagua D, Nievas-Soriano BJ. The problem of calculating the prevalence of sexual dysfunction: a meta-analysis attending gender. Sex Med Rev 2024 Mar 26;12(2):116-126. [doi: 10.1093/sxmrev/qead058] [Medline: 38336366]
16. Heidari M, Ghodusi M, Rezaei P, Kabirian Abyaneh S, Sureshjani EH, Sheikhi RA. Sexual function and factors affecting menopause: a systematic review. J Menopausal Med 2019 Apr;25(1):15-27. [doi: 10.6118/jmm.2019.25.1.15] [Medline: 31080785]
17. Hovén E, Fagerkvist K, Jahnukainen K, et al. Sexual dysfunction in young adult survivors of childhood cancer - a population-based study. Eur J Cancer 2021 Sep;154:147-156. [doi: 10.1016/j.ejca.2021.06.014] [Medline: 34273812]
18. Rahmani A, Afsharnia E, Fedotova J, et al. Sexual function and mood disorders among menopausal women: a systematic scoping review. J Sex Med 2022 Jul;19(7):1098-1115. [doi: 10.1016/j.jsxm.2022.03.614] [Medline: 35752457]
19. Calzo JP, Austin SB, Charlton BM, et al. Erectile dysfunction in a sample of sexually active young adult men from a U.S. cohort: demographic, metabolic and mental health correlates. J Urol 2021 Feb;205(2):539-544. [doi: 10.1097/JU.0000000000001367] [Medline: 32935616]
20. Bahrami Vazir E, Mohammad-Alizadeh-Charandabi S, Kamalifard M, Ghelichkhani F, Mohammadi A, Mirghafourvand M. The correlation between sexual dysfunction and intimate partner violence in young women during pregnancy. BMC Int Health Hum Rights 2020 Sep 14;20(1):24. [doi: 10.1186/s12914-020-00245-9] [Medline: 32928215]
21. Defeudis G, Mazzilli R, Tenuta M, et al. Erectile dysfunction and diabetes: a melting pot of circumstances and treatments. Diabetes Metab Res Rev 2022 Feb;38(2):e3494. [doi: 10.1002/dmrr.3494] [Medline: 34514697]
22. Jackson SE, Yang L, Koyanagi A, Stubbs B, Veronese N, Smith L. Declines in sexual activity and function predict incident health problems in older adults: prospective findings from the english longitudinal study of ageing. Arch Sex Behav 2020 Apr;49(3):929-940. [doi: 10.1007/s10508-019-1443-4] [Medline: 31432361]
23. Hamzehgardeshi Z, Sabetghadam S, Pourasghar M, Khani S, Moosazadeh M, Malary M. Prevalence and predictors of sexual distress in married reproductive-age women: A cross-sectional study from Iran. Health Sci Rep 2023 Sep;6(9):e1513. [doi: 10.1002/hsr2.1513] [Medline: 37655267]
24. Li L, Yu L. The influence of pension mode on the mental health of older adults—evidence from older adults in China. IJERPH 2021 Dec 23;19(1):119. [doi: 10.3390/ijerph19010119]
25. Fang EF, Xie C, Schenkel JA, et al. A research agenda for ageing in China in the 21st century (2nd edition): Focusing on basic and translational research, long-term care, policy and social networks. Ageing Res Rev 2020 Dec;64:101174. [doi: 10.1016/j.arr.2020.101174] [Medline: 32971255]
26. Wang B, Peng X, Liang B, et al. Sexual well-being among older adults in China (SWELL): protocol for a multicenter cross-sectional study. BMJ Open 2023 Jan 30;13(1):e067338. [doi: 10.1136/bmjopen-2022-067338] [Medline: 36717139]
27. Laumann EO, Gagnon JH, Michael RT, Michaels S. The Social Organization of Sexuality: Sexual Practices in the United States: University of Chicago press; 2000.
28. Dadvand P, Bartoll X, Basagaña X, et al. Green spaces and General Health: roles of mental health status, social support, and physical activity. Environ Int 2016 May;91:161-167. [doi: 10.1016/j.envint.2016.02.029]

29. Whittle HJ, Sheira LA, Wolfe WR, et al. Food insecurity is associated with anxiety, stress, and symptoms of posttraumatic stress disorder in a cohort of women with or at risk of HIV in the United States. J Nutr 2019 Aug 1;149(8):1393-1403. [doi: 10.1093/jn/nxz093] [Medline: 31127819]
30. Wang B, Peng X, Fu L, et al. Correlates of sexual lifestyles among older adults living with HIV in China: findings from the sexual well-being (SWELL) study. Infect Microb Dis 2023;5(3):145-151. [doi: 10.1097/IM9.0000000000000125]
31. Soysal P, Smith L. Sexual activity and successful aging. Eur Geriatr Med 2022 Dec;13(6):1235-1238. [doi: 10.1007/s41999-022-00694-6]
32. Lindau ST, Schumm LP, Laumann EO, Levinson W, O'Muircheartaigh CA, Waite LJ. A study of sexuality and health among older adults in the United States. N Engl J Med 2007 Aug 23;357(8):762-774. [doi: 10.1056/NEJMoa067423] [Medline: 17715410]
33. Field N, Mercer CH, Sonnenberg P, et al. Associations between health and sexual lifestyles in Britain: findings from the third national survey of sexual attitudes and lifestyles (Natsal-3). Lancet 2013 Nov 30;382(9907):1830-1844. [doi: 10.1016/S0140-6736(13)62222-9] [Medline: 24286788]
34. Lee DM, Nazroo J, O'Connor DB, Blake M, Pendleton N. Sexual health and well-being among older men and women in England: Findings from the English longitudinal study of ageing. Arch Sex Behav 2016 Jan;45(1):133-144. [doi: 10.1007/s10508-014-0465-1] [Medline: 25624001]
35. Kolotkin RL, Zunker C, Østbye T. Sexual functioning and obesity: a review. Obesity (Silver Spring) 2012 Dec;20(12):2325-2333. [doi: 10.1038/oby.2012.104] [Medline: 22522887]
36. Mollaioli D, Ciocca G, Limoncin E, et al. Lifestyles and sexuality in men and women: the gender perspective in sexual medicine. Reprod Biol Endocrinol 2020 Feb 17;18(1):10. [doi: 10.1186/s12958-019-0557-9] [Medline: 32066450]
37. Faubion SS, Fairbanks F, Kuhle CL, et al. Association between body mass index and female sexual dysfunction: a cross-sectional study from the data registry on experiences of aging, menopause, and sexuality. J Sex Med 2020 Oct;17(10):1971-1980. [doi: 10.1016/j.jsxm.2020.07.004] [Medline: 32771351]
38. Law S, Wan K, Yang W, Global Obesity C. Effects of bariatric surgery on sexual function and fertility: a narrative review. Obes Rev 2024 Aug;25(8):e13757. [doi: 10.1111/obr.13757] [Medline: 38689132]
39. Kumagai H, Yoshikawa T, Zempo-Miyaki A, et al. Vigorous physical activity is associated with regular aerobic exercise-induced increased serum testosterone levels in overweight/obese men. Horm Metab Res 2018 Jan;50(1):73-79. [doi: 10.1055/s-0043-117497] [Medline: 28934816]
40. McNabney SM, Gletsu-Miller N, Rowland DL. Sexual function and satisfaction in the context of obesity. Curr Diab Rep 2023 Nov;23(11):315-327. [doi: 10.1007/s11892-023-01523-1] [Medline: 37632680]
41. Santamaria A, Amighi A, Thomas M, Goradia R, Choy J, Hehemann MC. Effect of surgical, medical, and behavioral weight loss on hormonal and sexual function in men: a contemporary narrative review. Ther Adv Urol 2024;16:17562872241279648. [doi: 10.1177/17562872241279648] [Medline: 39285942]
42. Du Plessis SS, Cabler S, McAlister DA, Sabanegh E, Agarwal A. The effect of obesity on sperm disorders and male infertility. Nat Rev Urol 2010 Mar;7(3):153-161. [doi: 10.1038/nrurol.2010.6] [Medline: 20157305]
43. Smith L, Grabovac I, Veronese N, et al. Sleep quality, duration, and associated sexual function at older age: findings from the english longitudinal study of ageing. J Sex Med 2019 Mar;16(3):427-433. [doi: 10.1016/j.jsxm.2019.01.005] [Medline: 30773496]
44. Seftel AD. Re: Sleep quality, duration, and associated sexual function at older age: findings from the english longitudinal study of ageing. Journal of Urology 2019 Sep;202(3):443-443. [doi: 10.1097/01.JU.0000574448.61263.77]
45. Andersen ML, Tufik S. The association between sleep disturbances and erectile dysfunction during the COVID-19 pandemic. Sex Med Rev 2022 Apr;10(2):263-270. [doi: 10.1016/j.sxmr.2021.12.001] [Medline: 35042659]
46. Pigeon WR, Youngren W, Carr M, Bishop TM, Seehuus M. Relationship of insomnia to sexual function and sexual satisfaction: findings from the sleep and sex survey II. J Psychosom Res 2023 Dec;175:111534. [doi: 10.1016/j.jpsychores.2023.111534] [Medline: 37871507]
47. Zhang F, Xiong Y, Qin F, Yuan J. Short sleep duration and erectile dysfunction: a review of the literature. Nat Sci Sleep 2022;14:1945-1961. [doi: 10.2147/NSS.S375571] [Medline: 36325277]
48. Allen MS, Walter EE. Health-related lifestyle factors and sexual dysfunction: a meta-analysis of population-based research. J Sex Med 2018 Apr;15(4):458-475. [doi: 10.1016/j.jsxm.2018.02.008] [Medline: 29523476]
49. Laumann EO, Nicolosi A, Glasser DB, et al. Sexual problems among women and men aged 40-80 y: prevalence and correlates identified in the Global Study of Sexual Attitudes and Behaviors. Int J Impot Res 2005;17(1):39-57. [doi: 10.1038/sj.ijir.3901250] [Medline: 15215881]
50. Velurajah R, Brunckhorst O, Waqar M, McMullen I, Ahmed K. Erectile dysfunction in patients with anxiety disorders: a systematic review. Int J Impot Res 2022 Mar;34(2):177-186. [doi: 10.1038/s41443-020-00405-4] [Medline: 33603242]
51. Ciocca G, Di Lorenzo G, Comite G, et al. Sexual functioning in male patients suffering from depression and anxiety disorders. Eur psychiatr 2017 Apr;41(S1):s848-s849. [doi: 10.1016/j.eurpsy.2017.01.1681]
52. Yafi FA, Jenkins L, Albersen M, et al. Erectile dysfunction. Nat Rev Dis Primers 2016 Feb 4;2:16003. [doi: 10.1038/nrdp.2016.3] [Medline: 27188339]

53. Zhang H, Fan S, Yip PSF. Sexual dysfunction among reproductive-aged Chinese married women in Hong Kong: prevalence, risk factors, and associated consequences. J Sex Med 2015 Mar;12(3):738-745. [doi: 10.1111/jsm.12791] [Medline: 25470214]
54. van Lankveld J, Jacobs N, Thewissen V, Dewitte M, Verboon P. The associations of intimacy and sexuality in daily life: Temporal dynamics and gender effects within romantic relationships. J Soc Pers Relat 2018 May;35(4):557-576. [doi: 10.1177/0265407517743076] [Medline: 29899585]
55. Štulhofer A, Jurin T, Graham C, Janssen E, Træen B. Emotional intimacy and sexual well-being in aging European couples: a cross-cultural mediation analysis. Eur J Ageing 2020 Mar;17(1):43-54. [doi: 10.1007/s10433-019-00509-x] [Medline: 32158371]

Abbreviations

AIC: Akaike Information Criterion
aORs: adjusted odds ratios
GAD-7: generalized anxiety disorder-7
GSSAB: Global Study of Sexual Attitudes and Behaviours
PHQ-9: Patient Health Questionnaire
SQS: sleep quality scale
SRPs: Sexual response problems
SWELL: Sexual Well-being

Edited by G Seçkin; submitted 23.09.24; peer-reviewed by H Song, J Bian, T Tian; revised version received 17.01.25; accepted 23.02.25; published 01.05.25.

Please cite as:
Liang B, Xu C, Wang B, Li X, Peng X, Wang Y, Li H, Lu Y, Shen X, Ouyang L, Wu G, Yu M, Liu J, Meng X, Cai Y, Zou H
Sexual Response Problems and Their Correlates Among Older Adults From the Sexual Well-Being (SWELL) Study in China: Multicenter Cross-Sectional Study
JMIR Aging 2025;8:e66772
URL: https://aging.jmir.org/2025/1/e66772
doi:10.2196/66772

Older Adults' Perspectives on Participating in a Synchronous Online Exercise Program: Qualitative Study

Giulia Coletta[1], PhD; Kenneth S Noguchi[2], PhD; Kayla Beaudoin[1], BSc; Angelica McQuarrie[1], MSc; Ada Tang[2], PhD; Rebecca Ganann[3], PhD; Stuart M Phillips[1], PhD; Meridith Griffin[4,5], PhD

[1]Department of Kinesiology, McMaster University, Hamilton, ON, Canada

[2]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

[3]School of Nursing, McMaster University, Hamilton, ON, Canada

[4]Department of Health, Aging and Society, McMaster University, Kenneth Taylor Hall, 241, 1280 Main Street West, Hamilton, ON, Canada

[5]Gilbrea Centre for Studies in Aging, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Meridith Griffin, PhD
Department of Health, Aging and Society, McMaster University, Kenneth Taylor Hall, 241, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Older adults face several barriers to exercise participation, including transportation, lack of access, and poor weather conditions. Such barriers may influence whether older adults meet the Canadian 24-Hour Movement Guidelines. Recently, older adults have adopted technology for health care and are increasingly using digital health technologies to improve their access to care. Therefore, technology may be a valuable tool to reduce barriers to exercise and increase exercise participation rates within this population.

Objective: This study aimed to explore older adults' perceptions and experiences of exercise, in general, and specifically related to our synchronous online exercise program for community-dwelling older adults.

Methods: A total of 3 registered kinesiologists and 1 physiotherapist with experience working with older adults delivered an 8-week, thrice-weekly synchronous online group-based exercise program for older adults in 3 cohorts. The program focused on strength, balance, and aerobic activity. Following the program, a qualitative study with interpretive descriptive design was conducted to explore participants' perceptions and experiences. Participants were invited to take part in a 30-minute, one-on-one semistructured interview via Zoom with a research team member. Interview data were thematically analyzed to identify common themes.

Results: A total of 22 older adults (16 women, 6 men; mean age 70, SD 4 years) participated in interviews. Three themes were identified as follows: (1) health, exercise, and aging beliefs; (2) the pandemic interruption and impacts; and (3) synchronous online exercise programs attenuate barriers to exercise. Participants discussed their exercise beliefs and behaviors and their desire to safely and correctly participate in exercise. Older adults found that their physical activity was curtailed, routines disrupted, and access to in-person exercise programs revoked due to the pandemic. However, many suggested that our synchronous online exercise program was motivational and attenuated commonly reported environmental barriers to participation, such as transportation concerns (eg, time spent traveling, driving, and parking), accessibility and convenience by participating at a location of their choice, and removing travel-related concerns during poor weather conditions.

Conclusions: Given these reported experiences, we posit that synchronous online exercise programs may help motivate and maintain adherence to exercise programs for older adults. These findings may be leveraged to improve health outcomes in community-dwelling older adults.

Trial Registration: ClinicalTrials.gov NCT04627493; https://clinicaltrials.gov/study/NCT04627493

(JMIR Aging 2025;8:e66473) doi:10.2196/66473

KEYWORDS

exercise; older adults; qualitative study; qualitative; experience; attitude; opinion; perception; perspective; interview; internet; kinesiology; physiotherapy; synchronous; online; home-based; gerontology; geriatric; older; aging; physical activity

Introduction

Physical activity (PA) guidelines are not met by ≈85% of older adults in Canada [1,2]. The Canadian Society for Exercise Physiology recommends that older adults participate in 150 minutes of moderate-to-vigorous PA and 2 bouts of strengthening exercise each week [3], similar to the World Health Organization guidelines [4]. If PA guidelines are not met, older adults are at greater risk of adverse health outcomes, including falls, cardiovascular disease, mobility limitations, hospitalization, and mortality [4].

Older adults commonly report barriers to participating in PA and exercise. A systematic review by Spiteri et al [5] determined that environmental factors, resources, lack of assistance in managing change, and social influences are among the common barriers to PA participation for older adults aged 65 - 70 years. Baeret et al [6] reported similar barriers in their systematic review of 44 journal papers describing motivators and barriers to PA in older adults. Common environmental barriers for older adults include transportation, geographic accessibility, affordability, and weather [5-7]. Transportation barriers can be attributed to a lack of transportation to facilities, challenges navigating parking, and poor reliability of affordable transportation [5,6]. Lack of access to exercise facilities, the cost of exercise programming, and distance to programs contribute to environmental barriers [5,6]. Lastly, poor weather conditions reduce the likelihood of adults participating in recreational activities or attending recreation centers [7]. Among Canadians, 64% are less active in the winter compared with the summer [8]. Developing exercise programs to mitigate common environmental barriers may aid in improving older adults' abilities to participate in exercise and meet PA guidelines.

Recently, older adults have adopted technology for health care to improve access to care [9,10]. Technology may, therefore, be a valuable mode of delivery to reduce barriers to exercise and increase participation rates within this population [11]. Current evidence suggests that online exercise programs improve physical function and cognition in older adults, although these programs are exclusively asynchronous, web-based, or interactive exercise-based video games [12]. Older adults have reported not enjoying asynchronous exercise programs as they lack social connectivity, accountability, and adherence, making it difficult to stay motivated [12,13]. Recently, there has been a greater uptake of asynchronous and synchronous programming due to the COVID-19 pandemic [14-16]. A scoping review by Dagenais et al [16] described the nature and extent of the existing literature on online exercise programming among older adults and found 17 studies using asynchronous (9/17, 53%), synchronous (5/17, 29%), or both asynchronous and synchronous programming (3/17, 18%). For older adults, group-based synchronous programs have demonstrated improvements in muscle strength, balance, physical function, aerobic capacity, and quality of life [14]. Therefore, gaining insights into older adults' perceptions and experiences with synchronous programming is important.

We developed and delivered a synchronous online exercise program to help older adults engage in exercise while in-person programming was temporarily suspended [17]. The synchronous program was designed and delivered by registered kinesiologists and physiotherapists. The main aim of the parent trial was to investigate the preliminary effectiveness and feasibility of the synchronous online exercise program through a mixed method pilot randomized controlled trial (RCT) [17]. The aim of this paper was to explore older adult participants' perceptions and experiences of exercise, in general, and of our synchronous online community exercise program.

Methods

Design

This qualitative study was part of an 8-week community-based, pilot RCT [17]. An interpretive descriptive design explored community-dwelling older adults' experiences and perceptions of exercise, in general, and of the synchronous online exercise program [18]. Interpretive description offers a flexible approach to analyzing qualitative data within medical education research, as it can explore individuals' experiences while producing practical clinical outcomes [19]. The pilot RCT compared community-dwelling older adults aged 65 to 80 years participating in a synchronous online exercise program (ACTIVE) delivered on Zoom with a waitlist control group (CON). A total of 3 registered kinesiologists and 1 physiotherapist with experience working with older adults conducted the intervention online in 3 cohorts at the Physical Activity Centre of Excellence at McMaster University in Ontario, Canada. The intervention and all data collection, including the qualitative interviews, were conducted remotely via Zoom. Zoom meets McMaster University's privacy legislation. The study duration was from January 2021 to May 2022. The study was reported in accordance with the consolidated criteria for reporting qualitative research [20]. For the purpose of this paper, the qualitative findings alone will be reported.

Ethical Considerations

This study was approved by the Hamilton Integrated Research Ethics Board (#11429) and registered at ClinicalTrials.gov (NCT04627493). Participants were informed of the study content and procedures both verbally and in writing. Informed written and oral consent was obtained from all participants. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants did not receive compensation for their participation.

Participants

ACTIVE and CON group participants enrolled in our 8-week synchronous online program were invited via email to participate in an individual qualitative interview after completing the exercise program. The eligibility criteria for the full trial have been previously reported [17]. Briefly, they were healthy older community-dwelling women and men between the ages of 65 and 80 years, had access to the internet via a personal smartphone, tablet, or computer, and participated in ≤150 minutes of moderate to vigorous PA per week. To be eligible for the qualitative portion of the study, participants must have

remained enrolled in the study and needed to participate in the online exercise program. Interested participants were contacted within 2 weeks of the end of the synchronous online exercise program to schedule an interview time. With a purposeful subsample of interested participants from the pilot RCT, we aimed to conduct between 20 and 24 interviews, as data saturation typically occurs within this range [21].

Intervention

The 8-week intervention consisted of thrice-weekly 60-minute synchronous online exercise classes (totaling 24) delivered online via Zoom by health professionals (ie, registered kinesiologists and physiotherapists). Briefly, the group-based classes began with a 5-minute warm-up, 50 minutes of progressive strength (25 min), aerobic (20 min), and balance training (5 min), and a 5-minute cool-down [17]. The physiotherapist (n=1) and registered kinesiologists (n=3) were trained to work with older adults and adjusted exercise intensity weekly to meet participants' abilities. The instructors and research team members supervised and moderated each class to ensure participant safety and assist with technological difficulties. All interactions between participants and the research team were conducted remotely via Zoom; no pre-established relationships existed prior to the study.

Qualitative Data Collection

Semistructured, open-ended one-on-one interviews were conducted via Zoom by a study investigator (GC, female, PhD Candidate) with experience in qualitative research [22]. A qualitative interview guide (Multimedia Appendix 1) was developed with a qualitative methodology expert (MG) to generate discussion around participants' experiences and perceptions of exercise, in general, and with our synchronous online exercise program [22]. Questions were not provided to participants before the interview. However, participants were aware of the interviewer's personal goals (ie, PhD research) for conducting the research. Transcripts were transcribed verbatim from Zoom audio recordings, checked for accuracy, and not returned to participants.

Data Analysis

The interviews were thematically analyzed using the computer software Dedoose (version 9.0.17; SocioCultural Research Consultants, LLC) [23]. Two authors, GC and KSN, independently coded the first 2 transcripts and discussed initial patterns and themes to inform the preliminary coding book [24], which was shared and refined with input from the qualitative expert. The coding book was used to review subsequent transcripts and iteratively revised to reflect new concepts. Themes and subthemes were inductively developed and refined with the research team. At least 2 interviewers (GC, KSN, and MG) discussed any disagreements in transcription or content to reach a consensus [23,25].

Results

Participants' Characteristics

Of the 32 who participated in the pilot RCT, which determined the preliminary effectiveness of the synchronous online exercise program [17], 22 older women and men participated in semistructured interviews. Ten participants declined or did not respond to the invitation to participate in this study. Table 1 presents the descriptive characteristics of the sample. There were no major differences between groups for any variable. The majority of participants were women (25/32, 78%) and had completed postsecondary education (29/32, 91%). All interviews were approximately 30 minutes in duration by design.

Table . Descriptive characteristics for the ACTIVE, control group (CON), and no-interview participants (N=32).

	ACTIVE[a] (n=13)	CON[b] (n=9)	No interview (n=10)
Women, n (%)	9 (69)	7 (78)	9 (90)
Age (years), mean (SD)	70 (4)	70 (4)	74 (4)
Race, n (%)			
Caucasian	12 (92)	9 (100)	9 (90)
Indigenous	1 (8)	0 (0)	0 (0)
Asian	0 (0)	0 (0)	1 (10)
Highest level of education earned, n (%)			
Some school or high school diploma	2 (15)	1 (11)	1 (10)
Some college, vocational, or training school	1 (8)	1 (11)	0 (0)
College graduate or bachelor's degree	5 (38)	5 (56)	7 (70)
Postgraduate training or degree	5 (38)	2 (22)	2 (20)
Partner status, n (%)			
Never married	0 (0)	2 (22)	0 (0)
Divorced or separated	4 (31)	1 (11)	1 (10)
Widowed	3 (23)	1 (11)	3 (30)
Presently married	6 (46)	5 (56)	6 (60)
Living arrangement, n (%)			
Wife, husband, or partner	6 (46)	5 (56)	6 (60)
Children	1 (8)	0 (0)	0 (0)
Friends	0 (0)	1 (11)	0 (0)
Living alone	5 (38)	3 (33)	4 (40)
Other	1 (8)	0 (0)	0 (0)
Physically active prior to the pandemic, n (%)	7 (54)	5 (56)	6 (60)

[a]ACTIVE: exercise group.

[b]CON: waitlist control group.

Qualitative Findings

Overview

Our thematic analysis identified 3 main themes with respect to participants' perceptions of exercise in general and with the synchronous online exercise program: health, exercise, and aging beliefs; the pandemic interruption and impacts; and synchronous online exercise programs attenuate barriers to exercise (Textbox 1). Participants described various subthemes within these 3 main themes, including their exercise beliefs and behaviors and their desire to safely and correctly participate in exercise as an older adult. They also reflected upon the impact of the pandemic on their PA, routine, and lack of access to exercise programs. Participants discussed how the online synchronous format contributed to their motivation to exercise, reduced concerns about transportation, enabled participation from their desired training location, and allowed them to avoid poor weather conditions. Our findings provide greater insight into older adults' perceptions and experiences of exercise, in general, through the themes of health, exercise, and aging beliefs and the impact of the pandemic on older adults' PA. We also provide insight into the delivery of synchronous online community exercise programs and how common barriers to exercise may be attenuated.

Textbox 1. Qualitative themes and subthemes.

Health, exercise, and aging beliefs

- Exercise beliefs and behaviors.
- Desire to safely and correctly participate in exercise as an older adult.

The pandemic interruption and impacts

- Physical activity curtailed.
- Routines disrupted.
- Access to exercise programs revoked.

Synchronous online exercise programs attenuate barriers to exercise

- Synchronous programming is motivating.
- Removed concerns about transportation.
- Ease of participation from desired training location.
- Removed poor weather condition concerns.

Health, Exercise, and Aging Beliefs

Participants believed in the benefits of exercise and sought ways to become more physically active. Many had participated in PA throughout their lives, and they valued the importance of the health benefits of exercise. Three participants described their beliefs about the benefits of exercise and its importance to health and longevity.

> *Yeah, it's like exercise should be long-term or lifetime commitment… it can really benefit you through your own lifetime.* [CON1]
>
> *I have the attitude exercise is important for everybody, but it's especially important as we get older.* [CON2]
>
> *I need physical exercise. My body feels better.* [CON3]

Two participants discussed the reasons for seeking out opportunities to increase their PA.

> *If I joined the program at least it would give me a reason to get up and get started in exercise.* [ACTIVE4]
>
> *Since I've retired, I'm trying to find some kind of routine that would keep me active, and [help with] losing weight and trying to get healthy for longevity.* [ACTIVE5]

Participants expressed concern about their declining fitness levels. Three participants described their desire to continue participating in meaningful activities and believed that becoming more active would help them maintain or improve their health.

> *I was getting too unfit basically… If I think in terms of basic fitness, it would be functional fitness. Can I bend down to get something on the bottom shelf? Can I reach the top shelf? Can I get in and out of a bathtub?… Activities of daily living, but all of them, like the reaching bending stretching [exercise would help with this].* [CON6]
>
> *Because I want to be active in my older age. I don't want to be, you know, stuck in a chair and very limited in what I can do. I still want to travel… You know I want to be mobile.* [ACTIVE7]
>
> *What does motivate me is as I'm aging, I want to get healthier… I want to maybe see my grandkids get married, you know, things like that, so as I age, I get more motivation.* [ACTIVE8]

One participant described their motivation to become more physically active for health and vanity.

> *It's for health, and to be honest with you, it partly is vanity, too.* [CON2]

The participants in our study had prior beliefs regarding exercise and believed in the long-term benefits of exercise. Their motivation stemmed from their desire to improve their health and continue to participate in activities that are meaningful to them. While participants expressed a desire to be active, many disclosed that they lacked knowledge of how to exercise. One participant described that as an older woman, strength training is not an easily accessible exercise modality. Aerobic is traditionally the commonly prescribed type of exercise, but older adults, including women, are interested in learning about strength training.

> *I think as an older person and a woman, I'm less likely to do [strength training] on my own. I can go out and walk and do aerobic stuff, but it's the resistance work that I would be interested in.* [ACTIVE5]

Three participants described concerns over incorrect form, potentially leading to injury. Older adults discuss a lack of confidence, fear of injury, and desire for more supervision when exercising.

> *I think as we get older, certainly doing it the proper way even becomes more important than when you were young.* [ACTIVE9]

> *It's not like I didn't want to try and join some of those [traditional] gyms sometimes, but I didn't feel as confident about it because of some of the medical part [injuries].* [CON3]

> *But there's not a lot of supervision happening there [traditional gyms]… the opportunity to do things incorrectly kind of opens you for doing something wrong and hurting yourself.* [ACTIVE7]

Participants in this study desired to become more physically active and believed in the benefits of exercise. Their motivations for exercising varied from wanting to continue with their activities of daily living to longevity and vanity. However, our participants suggested that traditional gym settings do not offer adequate supervision for older adults as they learn to participate in exercise safely.

The Pandemic Interruption and Impacts

The COVID-19 pandemic interrupted participants' daily routines and limited access to PA during this study. The first cohort of participants began the study during the second wave (Beta variant) of the COVID-19 pandemic and completed the study during the third wave (Gamma variant). The second cohort of participants began the study during the third wave of the COVID-19 pandemic and completed the study as the fourth wave (Delta variant) began. The third cohort of the study started and completed the study during the fifth wave (Omicron variant) of the COVID-19 pandemic. Participants voiced several concerns about their own inactivity levels and how these were affected by COVID-19. For example, 3 participants spoke of the secondary consequences of the pandemic resulting in unhealthy habits.

> *I had joined the gym, and then COVID happened. So, then I've put on another 30 pounds, so now I'm 70 pounds over my optimum weight…. I don't want to risk it right now because of the COVID stuff. Yeah, and I'm not doing massages and not doing physiotherapy, so you know that's kind of hindered things too.* [ACTIVE10]

> *I got out first thing in the morning; it was a part of a routine during the pandemic. I lost all of that; consequently, I put on weight… when the pandemic hit, I literally turned into a couch potato.* [CON2]

> *I was really surprised at how de-conditioned I became over the course of the pandemic; I mean, it wasn't great before, but it became a lot worse during the pandemic.* [ACTIVE11]

The pandemic's secondary consequences left participants more sedentary during this period and disrupted their routines. For example, 2 participants described the challenge to keep track of weekdays.

> *COVID really slowed us down, and I didn't know what day of the week it was.* [CON1]

> *Well, yeah, until the program that you gave us, I was getting lost on what day of the week it was. At least I knew when it was Monday, Wednesday, Friday.* [ACTIVE10]

In addition to disruptions to participants' daily routines, the pandemic revoked access to exercise programs. The public health-mandated lockdowns resulted in canceled in-person exercise programs, further limiting access to PA and exercise. Prior to the pandemic, participants had been engaging in exercise programs and were members of local gyms. COVID-19 interrupted their participation in programming, and many found it difficult to motivate themselves to exercise.

> *… once COVID hit, other classes [exercise programs] closed, it was very difficult… it's almost impossible to continue [exercising]. And I'm not very good at self-motivating.* [CON1]

> *Okay, I go to the gym quite often, like maybe four times or five times a week and, of course, the gyms are closed [COVID], and I found it hard to get motivated to do stuff on my own.* [ACTIVE4]

> *I just started at the Mac gym in the Senior's Program in February, so I was just sort of getting into the three-times-a-week schedule when, of course, we stopped in March because of the pandemic.* [ACTIVE11]

The pandemic interrupted many participants' routines, including canceled exercise programs and memberships and curtailing PA.

Synchronous Online Exercise Programs Attenuate Barriers to Exercise

Participants reported that our synchronous online exercise program removed several common environmental barriers to exercise and motivated them to become more physically active. For example, one participant discussed enjoying exercising from home while still connecting to people in real time.

> *The fact that we got to do it at home… that was cool. I think, maybe something like that would probably motivate me again that there's interaction, it's real people, and it's done in real-time… I actually liked having it [online].* [ACTIVE8]

Another participant spoke about how synchronous programming, in general, allows them to continue to exercise when they are not comfortable attending in-person programming.

> *Yeah, if you have sniffles, you might still want to do the exercises, but you don't want to spread the disease…* [ACTIVE10]

A third participant described how our synchronous online exercise program was advantageous compared with in-person programming.

> *So, it's just me moving around in the room, not, you know, bumping into other people like in a small gym… We're always overheated. Here I have a fan blowing on me, and I don't have somebody complaining that wind is blowing on them because they don't like wind.* [ACTIVE10]

Our synchronous online exercise program motivated participants to exercise as it allows individuals to remain connected with others while participating safely and easily from their own homes. Participants described enjoying the ease of participating

in our program from home since it removed their concerns about transportation and reduced time spent traveling to and from the exercise program, which allowed more time to engage in other activities. Participants discussed their concerns about driving, traffic, and parking when attending in-person exercise programs at the university or recreational facilities. For example, one participant referred to the difficulty of driving and finding their way around campus.

> *… I didn't like driving to McMaster, and then you have to figure out where you're going… that was just a part that I don't care about… So, I like local these days, and the fact that I could do that on Zoom is great. I think it's the best thing ever.* [CON12]

Another participant spoke about not having to drive (or otherwise find their way) to the session and how it streamlined their exercise experience.

> *No, that was actually really, really good. So, because it's COVID, you don't have to worry about being in contact. But yes, I don't like the concept of having to drive somewhere to exercise. It just defeats the whole purpose… I liked that.* [ACTIVE5]

A third participant discussed the inefficiency of driving to the gym.

> *I always used to say if people just walk to the gym or bike to the gym they wouldn't even have to go in the gym. So yes, you've cut out that timepiece of having to go somewhere to do something.* [ACTIVE7]

Our online exercise program eliminated transportation concerns and was described as motivational as it allowed individuals to participate from their desired training location, including their homes and while on holiday. For example, one participant described the ease of connecting to our program while away from home.

> *I was out of town, okay? I ended up at my nephew's place, and he works the night shift. So, I didn't want to disturb him inside, so I went outside, and they had a nice strong Wi-Fi connection, and it worked out really good.* [CON1]

Another participant enjoyed the flexibility of our program, allowing them to maintain their routine while away from home.

> *We'll sign up for that. We can do it at the lake as well… With this type of activity, you can do it in your living room, you can do it out on the deck, and you can do it wherever you want. I think it's great that there's more flexibility, that you can take the program to where you are, rather than you go[ing] there.* [ACTIVE9]

Two participants discussed how offering the exercise program through a synchronous videoconference platform enabled participation due to the ease of access.

> *It [Zoom] made me kind of want to go on… how can you argue about getting up five minutes before it starts? And, yeah, it's really, it's fantastic.* [ACTIVE13]

> *Having it be virtual means that you don't have to worry about traffic and parking and driving.* [CON6]

The synchronous program allowed flexibility in how and where participants engaged with our program, easing participants' concerns about traveling in poor weather conditions. Older adults described their fears and worries about navigating poor weather conditions (eg, extreme heat, rain, snow, and ice) on foot or by car to attend exercise programs or to be physically active. One participant described enjoying not needing to leave the house during the winter months to exercise.

> *Yeah, absolutely. In the winter, too, you didn't have to, you know, shiver and so on.* [CON14]

Two participants spoke of the inconvenience of navigating the weather conditions when exercising.

> *To go warm up my car, put my boots on, then go to the gym, take my boots off... It just doesn't make sense to me.* [ACTIVE8]

> *It was helpful that I didn't have to go running and try to park somewhere and find something [parking] and go through a snowbank.* [CON3]

Another participant discussed barriers when leaving the house when weather conditions are poor and how exercising from home helped to ensure they could still exercise.

> *… if it's like a torrential rain or snowstorm or ice, I don't go out on ice. Yeah, you know, then you would do it from home.* [6ACTIVE10]

Participants described poor weather conditions as negatively impacting their PA levels. Overall, they enjoyed the ease of the videoconference platform and participating in a synchronous online exercise program.

Discussion

Principal Findings

This study provides novel insights into synchronous online exercise programs for older adults. Specifically, participants identified themes including health, exercise, and aging beliefs; the pandemic interruption and impacts; and synchronous online exercise programs attenuating barriers to exercise. Our findings suggested that participants in this study believed in the benefits of exercise and wanted to learn how to exercise safely. The pandemic curtailed their PA, disrupted their routines, and revoked access to exercise programs, leaving them seeking new approaches to exercise. We found that our synchronous online exercise program was motivational and may reduce commonly reported environmental barriers, including removing concerns about transportation (eg, time spent traveling, driving, and parking), improving access through ease of participation from desired training locations, and removing poor weather condition concerns. Therefore, participants expressed a preference for synchronous online delivery of exercise programs, which may be a valuable option for motivating older adults to become physically active while reducing common environmental barriers.

Our findings suggested that participants believed in the health benefits of exercise. Participants described motivation to

exercise due to their desire to be healthy, mobile, and participate in meaningful activities as they aged. Notably, the majority of participants in our study were previously active and had prior experience with exercise, which may influence their perceptions and experiences with exercise. Our findings are similar to findings from a qualitative study by Harrison et al [26] focused on understanding the barriers and motivators of 58 older adults (49/58, 84% females; 60‐85 years) residing in an urban community in Washington, DC, which suggested that the main benefits and reasons for exercising are prolonged life, more energy, and a stronger body. Other participants in our study reported their appearance as a motivator for exercise. While participants in our study discussed their motivations for exercising, they lacked confidence in their abilities to exercise safely, feared injury, and desired supervision. Participants described how traditional gyms lack adequate supervision and support for older adults learning how to exercise. Access to exercise programs specific to older adults is critical to help motivate individuals to exercise and maintain PA levels [5,6].

Our study was conducted during the COVID-19 pandemic, which has been shown to limit walking, biking, PA, and mobility while increasing sedentary behavior across all age groups [27]. As a result, participants discussed the interruption of the pandemic to their lives and its impact in our interviews. While some evidence suggests that older adults became more active in the pandemic [28], this was not the case for our participants, who reported experiencing changes to their PA levels and daily routines and found it difficult to motivate themselves to be physically active independently. Similarly, in a 2021 qualitative study by Petersen et al [29], 12 healthy Canadian adults (6/12, 50% females; 20‐70 years), participants described disruptions to their daily routines and changes in PA as a result of the pandemic. Our study builds upon these findings by providing insight into how the closure of gyms and canceled exercise programs impacted participants in our study, as many were active members before the pandemic. Limited access to exercise programs may lead to difficulty for older adults in motivating themselves to be physically active.

Social connectivity is a key motivator for older adults to maintain adherence to exercise programs [6,30]. Our findings suggest that synchronous online exercise programs are motivational for older adults as they can connect in real time with instructors and other participants, which is unique to synchronous programming compared with asynchronous programs as older adults identify the lack of real-time objective feedback [31]. However, some older adults in our study had previously noted that socializing may not be important when exercising, and they prefer to focus on exercising alone [17]. The synchronous program made our participants feel comfortable exercising when they could not attend in-person programs, and it was advantageous compared with in-person exercise as they did not need to share their space. Synchronous programs may be a useful exercise approach as many older adults are shifting towards more digital connections with family and friends beyond the pandemic, as seen in a sedentary behavior reduction intervention with older adults [32]. However, this may not be a generalizable finding, particularly for older adults in Canada with lower socioeconomic status and lower educational attainment, as there may be barriers to technological access and literacy [33,34]. Careful consideration is needed to develop accessible future interventions and programs when leveraging technology to deliver synchronous PA and exercise programs more broadly. Future work should consider using in-person recruitment to ensure greater diversity in studies [35].

In addition to technology barriers, other commonly cited barriers to participation in PA and exercise included environmental factors, access to resources, assistance in managing change, and social influences for older community-dwelling adults [5,6]. Our findings reflect the environmental barriers reported in the literature, including transportation, access to exercise facilities, and poor weather conditions. Participants in our study expressed concerns about driving and parking at fitness centers, but these transportation concerns were alleviated through our online exercise program, which enabled participants to focus on exercising. Participants described how the “Internet” was a good way to “deal” with transportation concerns, as many described the extra mental load of in-person programming. Previous research has highlighted the importance of transportation assistance in promoting the uptake of programs and maintenance of PA for community-dwelling older adults who are socially disadvantaged or experience disability [36-38]. Reducing transportation concerns may improve the accessibility of exercise programs and the ability for older adults to participate from their desired training location [38].

Barriers to exercise accessibility due to lack of time (family and work) and lack of facilities are other commonly reported environmental barriers [5,6]. In focus groups with recently retired older adults, they described feeling like they are relied upon by children, grandchildren, parents, and friends, making it difficult to prioritize structured PA [39]. For some, online programs provide convenience in terms of not needing transportation, allowing older adults to spend more time on other priorities [40]. Our participants commented on the ease of getting up 5 minutes before class and our program’s convenience. The frequency of our program, 3 times per week, was also feasible for older adults [17]. Our synchronous online exercise program improved participants’ access to exercise specific for older adults and allowed the flexibility of attending class from home while still remaining socially connected. It additionally created a psychologically safe place for participants compared with traditional gyms, which are dominated by younger adults and lack supervised training specifically for older adults. Creating convenient, safe, and synchronous programs is important for ongoing adherence [39].

Poor weather conditions (eg, winter, rain, extreme heat) in Canada are another environmental barrier that contributes to increased sedentary time [41]. Our findings suggest that ice, snow, rain, and extreme heat discourage exercise, particularly since most older adults must travel to a facility to participate in an exercise program. Participants suggested that online delivery may effectively reduce this environmental barrier as participants do not need to be concerned with “cleaning their cars off” or “walking on ice” to attend a training facility to exercise. In older Canadians, there is a decrease in PA levels during precipitation [42] and winter months [8]. Synchronous delivery of exercise may be a feasible approach to deliver programs broadly,

particularly to populations who may not access in-person programs easily. Future work should explore hybrid exercise programming models, including in-person and synchronous delivery.

Strengths and Limitations

Our study suggests that synchronous exercise programs are motivational for older adults as they can connect in real-time with instructors while eliminating common environmental barriers to in-person exercise programming. Interpretive design allowed us to address a complex experiential question while producing a practical outcome [19]. A potential limitation of this study is the limited diversity among our sample. Our study population was predominately Caucasian, with high education attainment, women between 65 and 80, and motivated to exercise. Thus, the findings do not represent the perceptions and experiences of structurally marginalized populations, those experiencing lower socioeconomic status, men, older than 80 years, and those with low levels of education attainment or motivation to support behavioral change. We did not include nonparticipants in the program (n=1) in the interviews. Including nonparticipants may have allowed us to elucidate further barriers and health inequities for those unable to complete the program. Although the study lead provided technological support and assistance through Zoom or a phone call, including turning on cameras, muting and unmuting, and connecting speakers, all participants had good technology literacy and digital access. Future studies should consider including individuals in different populations with varying degrees of technological literacy. Additionally, conducting the synchronous online exercise program outside of the COVID-19 context would provide valuable insight into the long-term feasibility of this program. Trustworthiness may not have been achieved as the transcripts and codes were not checked and confirmed by the participants in this study [25]; however, triangulation across analysts strengthened this aspect of rigor.

Conclusion

The delivery of exercise programs using synchronous online classes may help older adults meet PA guidelines. Older adults report experiencing environmental barriers, including transportation, lack of access, and poor weather conditions when exercising. Our findings suggest that synchronous online exercise programs may serve as an approach to mitigate these environmental barriers and motivate older adults while keeping them socially connected. Considerations for designing exercise programs for older adults include delivery of the program by exercise and health professionals, synchronous programming to maintain social connectivity, and reducing environmental barriers such as transportation and weather concerns. Community programs may consider implementing synchronous online exercise programs as part of their recreational programming for older adults to increase engagement and reduce accessibility barriers. Future work should focus on leveraging synchronous exercise programs in community programs to engage older adults in PA and exercise and explore hybrid (synchronous and in-person) options.

Acknowledgments

We would like to acknowledge and thank the Physical Activity Centre of Excellence (PACE) registered kinesiologists and physiotherapists for their assistance in designing and delivering our online exercise program.

Conflicts of Interest

SMP reports grants or research contracts from the US National Dairy Council, Canadian Institutes for Health Research, Dairy Farmers of Canada, Roquette Freres, Ontario Centre of Innovation, Nestle Health Sciences, Myos, National Science and Engineering Research Council, and the US National Institutes of Health during the conduct of the study; personal fees from Nestle Health Sciences and nonfinancial support from Enhanced Recovery, outside the submitted work. SMP has patents licensed to Exerkine but reports no financial gains from patents or related work.

Multimedia Appendix 1
Semistructured interview questions.
[DOCX File, 18 KB - aging_v8i1e66473_app1.docx]

References

1. Bennie JA, De Cocker K, Teychenne MJ, Brown WJ, Biddle SJH. The epidemiology of aerobic physical activity and muscle-strengthening activity guideline adherence among 383,928 U.S. adults. Int J Behav Nutr Phys Act 2019 Apr 18;16(1):34. [doi: 10.1186/s12966-019-0797-2] [Medline: 30999896]
2. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Health Rep 2011 Mar;22(1):7-14. [Medline: 21510585]
3. Ross R, Chaput JP, Giangregorio LM, et al. Canadian 24-Hour Movement Guidelines for adults aged 18-64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl. 2):S57-S102. [doi: 10.1139/apnm-2020-0467] [Medline: 33054332]
4. Bull FC, Al-Ansari SS, Biddle S, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462. [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]

5. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
6. Baert V, Gorus E, Mets T, Geerts C, Bautmans I. Motivators and barriers for physical activity in the oldest old: a systematic review. Ageing Res Rev 2011 Sep;10(4):464-474. [doi: 10.1016/j.arr.2011.04.001] [Medline: 21570493]
7. Chan CB, Ryan DA. Assessing the effects of weather conditions on physical activity participation using objective measures. Int J Environ Res Public Health 2009 Oct;6(10):2639-2654. [doi: 10.3390/ijerph6102639] [Medline: 20054460]
8. Merchant AT, Dehghan M, Akhtar-Danesh N. Seasonal variation in leisure-time physical activity among Canadians. Can J Public Health 2007;98(3):203-208. [doi: 10.1007/BF03403713] [Medline: 17626385]
9. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023 May 16;6:e44564. [doi: 10.2196/44564] [Medline: 37191976]
10. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
11. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
12. Raquel Costa-Brito A, Bovolini A, Rúa-Alonso M, et al. Home-based exercise interventions delivered by technology in older adults: a scoping review of technological tools usage. Int J Med Inform 2024 Jan;181:105287. [doi: 10.1016/j.ijmedinf.2023.105287] [Medline: 37972483]
13. Goethals L, Barth N, Guyot J, Hupin D, Celarier T, Bongue B. Impact of home quarantine on physical activity among older adults living at home during the COVID-19 pandemic: qualitative interview study. JMIR Aging 2020 May 7;3(1):e19007. [doi: 10.2196/19007] [Medline: 32356777]
14. Fuentes Diaz MF, Leadbetter B, Pitre V, Nowell S, Sénéchal M, Bouchard DR. Synchronous group-based online exercise programs for older adults living in the community: a scoping review. J Aging Phys Act 2024 Dec 1;32(6):703-717. [doi: 10.1123/japa.2023-0214] [Medline: 38823794]
15. Peng R, Chang J, Du Y, et al. Older adults' perceptions and experiences of engaging in web- and mobile-based physical activity interventions: a systematic review and qualitative meta-synthesis. Geriatr Nurs 2024;59(630-8):630-638. [doi: 10.1016/j.gerinurse.2024.08.025] [Medline: 39197354]
16. Dagenais M, Parker O, Galway S, Gammage K. Online exercise programming among older adults: a scoping review. J Aging Phys Act 2023 Apr 1;31(2):289-302. [doi: 10.1123/japa.2021-0417] [Medline: 36087933]
17. Coletta G, Noguchi KS, Beaudoin KD, et al. A live online exercise program for older adults improves depression and life-space mobility: a mixed-methods pilot randomized controlled trial. PLoS One 2024;19(11):e0312992. [doi: 10.1371/journal.pone.0312992] [Medline: 39527532]
18. Thorne S. Interpretive Description: Qualitative Research for Applied Practice: Taylor & Francis; 2016.
19. Thompson Burdine J, Thorne S, Sandhu G. Interpretive description: a flexible qualitative methodology for medical education research. Med Educ 2021 Mar;55(3):336-343. [doi: 10.1111/medu.14380] [Medline: 32967042]
20. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Martínez-Mesa J, González-Chica DA, Duquia RP, Bonamigo RR, Bastos JL. Sampling: how to select participants in my research study? An Bras Dermatol 2016;91(3):326-330. [doi: 10.1590/abd1806-4841.20165254] [Medline: 27438200]
22. Adams WC. Conducting semi-structured interviews. In: Handbook of Practical Program Evaluation: John Wiley & Sons; 2015:492-505. [doi: 10.1002/9781119171386]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Methods Research: SAGE Publications; 2017.
25. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1). [doi: 10.1177/1609406917733847]
26. Harrison EA, Lord LM, Asongwed E, et al. Perceptions, opinions, beliefs, and attitudes about physical activity and exercise in urban-community-residing older adults. J Prim Care Community Health 2020;11:2150132720924137. [doi: 10.1177/2150132720924137] [Medline: 32468912]
27. Park AH, Zhong S, Yang H, Jeong J, Lee C. Impact of COVID-19 on physical activity: a rapid review. J Glob Health 2022;12(5003):05003. [doi: 10.7189/jogh.12.05003] [Medline: 35493780]
28. Colley RC, Saunders TJ. The ongoing impact of the COVID-19 pandemic on physical activity and screen time among Canadian adults. Health Rep 2023 Oct 18;34(10):13-23. [doi: 10.25318/82-003-x202301000002-eng] [Medline: 37856572]
29. Petersen JA, Naish C, Ghoneim D, Cabaj JL, Doyle-Baker PK, McCormack GR. Impact of the COVID-19 pandemic on physical activity and sedentary behaviour: a qualitative study in a Canadian city. Int J Environ Res Public Health 2021 Apr 22;18(9):4441. [doi: 10.3390/ijerph18094441] [Medline: 33922094]

30. Killingback C, Tsofliou F, Clark C. Older people's adherence to community-based group exercise programmes: a multiple-case study. BMC Public Health 2017 Jan 25;17(1):115. [doi: 10.1186/s12889-017-4049-6] [Medline: 28122532]
31. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022 Mar 25;22(1):248. [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
32. Greenwood-Hickman MA, Dahlquist J, Cooper J, et al. "They're Going to Zoom It": a qualitative investigation of impacts and coping strategies during the COVID-19 pandemic among older adults. Front Public Health 2021;9(679976):679976. [doi: 10.3389/fpubh.2021.679976] [Medline: 34095079]
33. Arcury TA, Sandberg JC, Melius KP, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150. [doi: 10.1177/0733464818807468] [Medline: 30353776]
34. Davidson J, Schimmele C. Evolving internet use among canadian seniors. Statistics Canada. 2019 Jul 7. URL: https://www150.statcan.gc.ca/n1/pub/11f0019m/11f0019m2019015-eng.htm [accessed 2025-03-27]
35. Carter CR, Maki J, Ackermann N, Waters EA. Inclusive recruitment strategies to maximize sociodemographic diversity among participants: a St. Louis case study. MDM Policy Pract 2023;8(1):23814683231183646. [doi: 10.1177/23814683231183646] [Medline: 37440792]
36. Nau T, Nolan G, Smith BJ. Enhancing engagement with socially disadvantaged older people in organized physical activity programs. Int Q Community Health Educ 2019 Jul;39(4):257-267. [doi: 10.1177/0272684X18821301] [Medline: 30638424]
37. Phoenix C, Griffin M, Smith B. Physical activity among older people with sight loss: a qualitative research study to inform policy and practice. Public Health 2015 Feb;129(2):124-130. [doi: 10.1016/j.puhe.2014.10.001] [Medline: 25687710]
38. Maula A, LaFond N, Orton E, et al. Use it or lose it: a qualitative study of the maintenance of physical activity in older adults. BMC Geriatr 2019 Dec 12;19(1):349. [doi: 10.1186/s12877-019-1366-x] [Medline: 31830900]
39. Van Dyck D, Mertens L, Cardon G, De Cocker K, De Bourdeaudhuij I. Opinions toward physical activity, sedentary behavior, and interventions to stimulate active living during early retirement: a qualitative study in recently retired adults. J Aging Phys Act 2017 Apr;25(2):277-286. [doi: 10.1123/japa.2015-0295] [Medline: 27759483]
40. Islam MS, Frazier MC, Harden SM, Lim S. Barriers and benefits of online group exercise programs for older adults. J Appl Gerontol 2024 Oct;43(10):1397-1407. [doi: 10.1177/07334648241240599] [Medline: 38536888]
41. Turrisi TB, Bittel KM, West AB, et al. Seasons, weather, and device-measured movement behaviors: a scoping review from 2006 to 2020. Int J Behav Nutr Phys Act 2021 Feb 4;18(1):24. [doi: 10.1186/s12966-021-01091-1] [Medline: 33541375]
42. Hoppmann CA, Lee JCM, Ziegelmann JP, Graf P, Khan KM, Ashe MC. Precipitation and physical activity in older adults: the moderating role of functional mobility and physical activity intentions. J Gerontol B Psychol Sci Soc Sci 2017 Sep 1;72(5):792-800. [doi: 10.1093/geronb/gbv107] [Medline: 26707498]

Abbreviations

CON: control group
PA: physical activity
RCT: randomized controlled trial

Edited by M O'Connell; submitted 14.09.24; peer-reviewed by B Leadbetter, IJ Liang, S Harden; revised version received 31.12.24; accepted 10.03.25; published 03.04.25.

Please cite as:
Coletta G, Noguchi KS, Beaudoin K, McQuarrie A, Tang A, Ganann R, Phillips SM, Griffin M
Older Adults' Perspectives on Participating in a Synchronous Online Exercise Program: Qualitative Study
JMIR Aging 2025;8:e66473
URL: https://aging.jmir.org/2025/1/e66473
doi:10.2196/66473

Effect of Physical Exercise on Telomere Length: Umbrella Review and Meta-Analysis

Juan Luis Sánchez-González[1], PhD; Juan Luis Sánchez-Rodríguez[2], PhD; Rogelio González-Sarmiento[3], PhD; Víctor Navarro-López[4], PhD; Raúl Juárez-Vela[5], PhD; Jesús Pérez[6,7,8,9], PhD; Javier Martín-Vallejo[10], PhD

[1]Faculty of Nursing and Physiotherapy, Department of Physiotherapy, University of Salamanca, Salamanca, Spain

[10]Faculty of Medicine, Department of Statistics, University of Salamanca, Salamanca, Spain

[2]Faculty of Psychology, Department of Basic Psychology, Psychobiology and Methodology, University of Salamanca, Salamanca, Spain

[3]Faculty of Medicine, Department of Medicine, University of Salamanca, Salamanca, Spain

[4]Faculty of Health Sciences, Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Universidad Rey Juan Carlos, Madrid, Spain

[5]Faculty of Health Sciences, Department of Nursing, University of La Rioja, Logroño, Spain

[6]Faculty of Medicine, Department of Psychiatry, University of Salamanca, Avenida Donantes de Sangre s/n, Salamanca, Spain

[7]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

[8]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, United Kingdom

[9]Norwich Medical School, University of East Anglia, Norwich, United Kingdom

Corresponding Author:
Jesús Pérez, PhD
Faculty of Medicine, Department of Psychiatry, University of Salamanca, Avenida Donantes de Sangre s/n, Salamanca, Spain

Abstract

Background: Telomere length (TL) is a marker of cellular health and aging. Physical exercise has been associated with longer telomeres and, therefore, healthier aging. However, results supporting such effects vary across studies. Our aim was to synthesize existing evidence on the effect of different modalities and durations of physical exercise on TL.

Objective: The aim of this study was to explore the needs and expectations of individuals with physical disabilities and their interventionists for the use of a virtual reality physical activity platform in a community organization.

Methods: We performed an umbrella review and meta-analysis. Data sources included PubMed, Embase, Web of Science, Cochrane Library, and Scopus. We selected systematic reviews and meta-analyses of randomized and nonrandomized controlled clinical trials evaluating the effect of physical exercise on TL.

Results: Our literature search retrieved 12 eligible systematic reviews, 5 of which included meta-analyses. We identified 22 distinct primary studies to estimate the overall effect size of physical exercise on TL. The overall effect size was 0.28 (95% CI 0.118-0.439), with a heterogeneity test value Q of 43.08 (P=.003) and I^2 coefficient of 51%. The number of weeks of intervention explained part of this heterogeneity (Q_B=8.25; P=.004), with higher effect sizes found in studies with an intervention of less than 30 weeks. Exercise modality explained additional heterogeneity within this subgroup (Q_B=10.28, P=.02). The effect sizes were small for aerobic exercise and endurance training, and moderate for high-intensity interval training.

Conclusions: Our umbrella review and meta-analysis detected a small-moderate positive effect of physical exercise on TL, which seems to be influenced by the duration and type of physical exercise. High quality studies looking into the impact of standardized, evidence-based physical exercise programs on TL are still warranted.

Trial Registration: PROSPERO CRD42024500736; https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=500736

(JMIR Aging 2025;8:e64539) doi:10.2196/64539

KEYWORDS

aging; chromosome; exercise; meta-analysis; telomere; telomerase; genes; genome; DNA

Introduction

Life expectancy has increased worldwide over the last century. The United Nations World Population Prospects [1] estimates that the world's population aged 65 years or older will rise by 16% in 2050, doubling the number of children younger than 5 years. This prospect will lead to a significant increase in age-related illnesses, such as cancer, dementia, or cardiovascular diseases [2], and is prompting research into the primary

mechanisms of aging and, subsequently, strategies to promote a healthy late-life.

López-Otín et al [3] have recently postulated 12 basic biological mechanisms of aging: epigenetic alterations, loss of proteostasis, disabled macroautophagy, deregulated nutrient-sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, altered intercellular communication, chronic inflammation, dysbiosis, genomic instability, and telomere attrition. Indeed, one of the main theories that has attempted to explain aging relates to telomere length (TL) and the role of telomerase, an enzyme responsible for maintaining and elongating telomeres. Telomeres are essential nucleoprotein structures located at the termini of eukaryotic chromosomes that play pivotal roles in safeguarding genomic integrity and regulating processes such as tumor suppression and aging [4]. Typically composed of a repetitive guanine-rich sequence extending from the chromosome end in the 5’ to 3’ orientation, paired with a complementary cytidine-rich strand [5], telomeres exhibit variations in sequence among species, yet share a common repetitive pattern across organisms. Measuring around 15 kilobases at birth in human somatic cells, telomeres undergo gradual attrition, with approximately 25 to 200 bases lost from their ends during each cell division in the absence of telomerase [6]. Upon reaching a critical length, telomeres signal cell cycle arrest, leading to cellular senescence and eventual demise [7]. Telomeres serve to shield chromosome ends from degradation and fusion, thereby upholding genomic stability [8,9].

There are different demographic factors that influence TL, such as genetics that may explain the high heritability of TL [10,11]; sex, with longer telomeres found in adult females [12]; or ethnicity, with the White community usually having longer telomeres than Black or Hispanic communities [13]. Also, stress levels have been associated with shorter TL, which may be due to oxidative stress and reduced telomerase enzyme activity [14,15]. Obesity [16], alcohol consumption [17], and smoking [16,18] are also factors that negatively influence TL.

Notably, longer telomeres have been found in people who exercise regularly [19] or with higher levels of daily physical activity [20,21]. Indeed, one of the most studied interventions regarding TL is physical exercise. Recent systematic reviews and meta-analyses [22-26] have investigated the effect of physical exercise on the TL of clinical and nonclinical samples, showing some positive, but still inconclusive, results. Aerobic exercise, such as running or swimming, has been consistently associated with longer telomeres. Denham et al [27] conducted a study that found individuals who regularly engage in aerobic exercise show greater telomerase activity, the enzyme that helps maintain TL, compared with those who lead a sedentary lifestyle. In addition, Ludlow et al [28] demonstrated that resistance training, which includes exercises like weight lifting, may also have beneficial effects on TL, possibly through mechanisms involving the reduction of oxidative stress and inflammation. In addition, more recent research has indicated that high-intensity training, such as high-intensity interval training (HIIT), may be more effective in preserving TL compared with low- to moderate-intensity exercises [29]. This type of exercise can induce more robust adaptive responses at the cellular level, including increased expression of genes related to longevity and telomere protection [3].

The aim of this umbrella review was to synthesize existing evidence on the effect of different modalities and durations of physical exercise on TL to inform the implementation of evidence-based physical exercise programs or recommendations to add healthier years into longer lives.

Methods

Overview

This umbrella review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [30], and it is registered in the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD42024500736).

Inclusion Criteria

The inclusion criteria and description of studies in this review followed the PICOS (Population, Intervention, Control, Outcomes, and Study design) framework for reviews [31].

Study Design

We selected systematic reviews, with or without meta-analyses, of randomized and nonrandomized controlled clinical trials, excluding nonexperimental designs, such as observational studies. No restrictions were applied on the basis of any particular language, following current international recommendations [32].

Population

Study participants were people with or without a clinical condition. The systematic reviews had to explicitly state that they included humans in their analyses. Therefore, studies involving experimental animals, such as rodents, were excluded.

Intervention and Control

We selected all systematic reviews and meta-analyses comparing the effects of physical exercise versus other intervention, or none, on TL. In addition, we further divided the results according to different types of physical exercise. If any reviews included primary studies combining different types of exercise, they were classified as “combined.”

Outcome Measure

The outcome measure was TL-related calculations.

Search Strategy

We performed a literature search for scientific articles published until January 19, 2024, using the following databases: PubMed (Medline), PEDro, Embase, Web of Science, and Cochrane Library. Multimedia Appendix 1 provides the search strategies, tailored for each database. A total of 2 independent reviewers (JLS-G and JLS-R) conducted the search using the same methodology. Any discrepancies during this process were resolved through consensus including a third reviewer (VN-L). In addition, we manually examined the reference sections of the original studies, and, if needed, contacted the authors for additional information.

Selection Criteria and Data Extraction

A total of 2 independent reviewers (JLS-G and JLS-R) initiated a screening process to evaluate relevance of the systematic reviews and meta-analyses. The initial screening was based on the title, abstract, and keywords for each review. When there was no consensus, or if the abstracts provided insufficient information, the full text was examined. During the second screening phase, the full text was evaluated to confirm inclusion criteria. Data presented in the results section were extracted using a protocol to ensure retrieval of the most relevant information for each study. The sample sizes, type of exercise, duration, means, and SDs of the telomere size data were extracted for each of the studies included in this review. During the analysis, we checked whether the results were repeated in different reviews or meta-analyses; if the effect sizes coincided, only one of the effect sizes was chosen; if the effect sizes did not match, the primary study was selected.

Methodological Quality Assessment

A total of 2 independent reviewers (JLSG and VNL) evaluated the methodological quality of the systematic reviews and meta-analyses, using the modified quality assessment scale for systematic reviews (AMSTAR-2 [A Measurement Tool to Assess Systematic Reviews]). AMSTAR-2 [33] is a questionnaire comprised of 16 domains, with simple categorical options: "yes," when the result is positive; "no," when the standard is not met or there is not enough information to answer it; and "partial yes," when there was partial adherence to the standard. In addition, we calculated the kappa coefficient (κ) and percentage (%) agreement scores to assess reliability before reaching consensus. Interrater reliability was estimated using κ, where κ>0.7 indicates a high level of agreement between reviewers, κ of 0.5 - 0.7 indicates a moderate level of agreement, and κ<0.5 suggests a low level of agreement [34].

Risk of Bias Assessment

We evaluated the risk of bias using the Risk Of Bias in Systematic reviews (ROBIS) assessment, which includes three domains: (1) relevance of assessment (optional); (2) identification of concerns with the review process through 4 domains related to study eligibility criteria, identification and selection of studies, data collection and study appraisal, and synthesis and findings; and (3) judgment on the risk of bias. The ROBIS tool includes signaling questions to assess specific domains and to guide the judgment of the systematic review's risk of bias, with responses categorized as "yes," "probably yes," "probably no," "no," or "no information." The risk of bias is then categorized as "low," "high," or "unclear" [35].

A total of 2 independent reviewers (JLS-G and JLS-R) assessed the risk of bias in the selected studies. In addition, we computed the kappa coefficient (κ) and percentage (%) agreement scores to evaluate reliability before reaching consensus.

Grading of Evidence

The Physical Activity Guidelines Advisory Committee Grading Criteria (PAGAC) were used to evaluate grading of evidence. The criteria used for assessing the quality of evidence included (1) applicability of the study sample, exposures, and outcomes to the research question; (2) generalizability to the population of interest; (3) risk of bias/study limitations; (4) quantity and consistency of findings across studies; and (5) magnitude and precision of the effect. Based on this information, final evidence grades and conclusion statements for each research question were formulated [36].

Overlap of Primary Studies

The overlap analysis of primary studies among the systematic reviews was performed with the Graphical Representation of Overlap for OVErviews (GROOVE) tool [37]. Using a matrix of evidence, GROOVE establishes the number of primary studies and systematic reviews, the absolute number of overlapped and nonoverlapped primary studies, and an overall corrected covered area (CCA) assessment. GROOVE also offers a detailed CCA assessment for each possible pair of systematic reviews (or "nodes"), with a graphical and easy-to-read representation of these results. In addition, it provides an optional function that incorporates the structural missingness within the matrix.

Data Synthesis and Analysis

Given the high variability of designs, patients and endpoints across studies, results were integrated following a random effects model. The Hedges unbiased standardized mean difference was used to determine effect sizes. Heterogeneity was measured using Q-tests, I^2 coefficients, and prediction intervals [38]. Metaregression and metapartition [39] were used to explain heterogeneity. These techniques involve partitioning sum of squares of the homogeneity test into 2 components, Q_B and Q_W, to account for quantitative or qualitative variables. If the Q_E value was large and statistically significant compared with the value of Q_W, it was considered that this variable explained part of the heterogeneity found in the integration of results across studies. If the variable was qualitative, subgroups of studies, which differed in the effect size and also showed less heterogeneity, were created. The procedure was repeated for subgroups that still presented large heterogeneity for other variables. Publication bias was assessed using the contour-enhanced funnel plot, Egger test, Doi plot, and LFK index procedures. 95% CIs were calculated for effect sizes. The significance level was set at 5%. The analyses were carried out with the *meta* v7.0, *metafor* v4.6 and *metasens* 1.5 - 2 libraries of the R statistics software (version 4.4.0; R Foundation for Statistical Computing).

Ethical Considerations

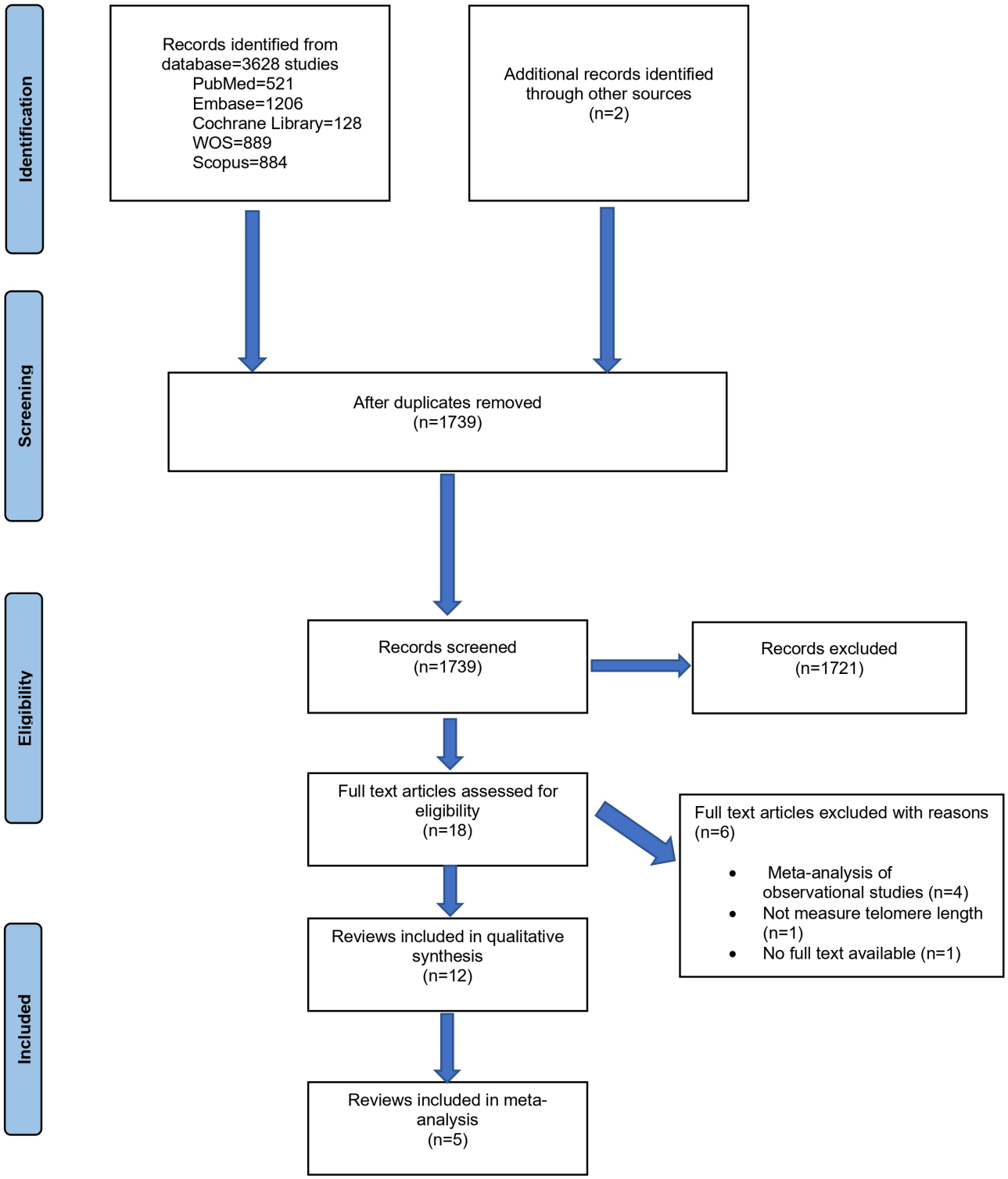
As this is an umbrella review, it is not necessary to have the approval of the ethics committee.

Results

Study Selection

The initial literature search revealed 3628 records. In addition, 2 more were retrieved manually from the references. A total of 12 systematic reviews were eligible for qualitative synthesis, and 5 of them allowed a meta-analysis. Figure 1 shows the study screening strategy.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Characteristics of the Systematic Reviews

Table 1 lists the characteristics of the systematic reviews and meta-analyses (type of study, design, sample, intervention, lab technique to measure TL, risk of bias, evaluation of quality, and conclusions).

Table . Characteristics of the systematic reviews.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Song et al [26]	Systematic review and meta-analysis	7 (RCCTs)	0	939	Healthy (3), breast cancer (2), polycystic ovary syndrome (1), and obese (1)	Aerobic exercise (4) and endurance training (3)	48 (2), 24 (3), and 16 (2)	qPCR[d]	Yes	No	The type and duration of exercise for positive improvement in TL is aerobic exercise for more than 6 months.
Buttet et al [22]	Systematic review and meta-analysis	13 (RCCTs) and 3 (CCTs)	5	908	Healthy (5), obese (1), and myocardial infarction (1)	Endurance training (7), strength (1), and HIIT[e] (1)	48 (2), 32 (1), and 24 (4)	qPCR	No	No	A lifestyle intervention with physical activity + diet can increase TL, independently of population characteristics or baseline TL.
Sánchez-González et al [19]	Systematic review and meta-analysis	8 (RCCTs) and 1 (CT)	0	1320	Healthy (9)	Endurance training (4), aerobic exercise (4), HIIT (3), and aerobic + endurance (1)	8 (1), 16 (1), 24 (4), and 52 (3)	qPCR	Yes	Yes	The findings suggest that HIIT seems to have a positive effect on TL compared with other types of exercise such as endurance training or aerobic exercise in healthy population.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Valente et al [25]	Systematic review and meta-analysis	6 (RCCTs)	16 (case-control studies), 2 (prospective cohort), 3 (cross-sectional), and 3 (retrospective cohort)	612	Healthy (5) and obese (1)	Endurance training (5), strength (1), aerobic exercise (1), and HIIT (1)	56 (1), 48 (1), and 24 (4)	qPCR	Yes	No	There is very low certainty that physically active individuals have longer telomeres with a moderate effect, but this effect is probably overestimated.
Denham et al [23]	Systematic review and meta-analysis	12 (NR[f])	0	487	Healthy (6) and chronic fatigue (1)	HIIT (1), aerobic exercise (3), Pilates training (1), and endurance training (2)	52 (1), 24 (2), 22 (1), 8 (2), and 1 (1)	qPCR	Yes	No	Exercise training as an inexpensive lifestyle factor that increases telomerase reverse transcriptase expression and telomerase activity. Regular exercise training could attenuate telomere attrition through a telomerase-dependent mechanism and ultimately extend healthspan longevity.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Barragán et al [40]	Systematic review	12 (NR)	53 (cross-sectional studies), 13 (case-control), and 9 (longitudinal)	1.056	Healthy (12)	Endurance training (5), combined training (1), aerobic exercise (6), NR (1), and HIIT (1)	56 (1), 52 (2), 24 (3), 8 (2), 1 (2), and NR (2)	qPCR	No	Yes	Although fewer sedentary activities, optimal sleep habits, and non- or ex-smoker status have been associated with less telomere shortening, several methodological issues were detected, including the need for more targeted interventions and standardized protocols to better understand how physical activity and sleep can impact TL and aging.
Schellnegger et al [41]	Systematic review	8 (RCCTs) and 7 (CCTs)	27 (observational studies)	1.700	Healthy (13) and obese (2)	Aerobic exercise (8), endurance training (8), and HIIT (1)	1 (6), 6 (1), 8 (1), 12 (1), 24 (5), and 52 (2)	qPCR	No	No	Physical activity with regular aerobic training of moderate to vigorous intensity appears to help preserve TL.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Prathap et al [42]	Systematic review	2 (RCCTs): 1 in rodents	3 (literature review)	151	Breast cancer (1)	Aerobic exercise (1)	24 (1)	qPCR	No	No	Based on the evidence collected it can be suggested that chronic moderate intensity aerobic exercise in a lifelong practice shows beneficial effects in a dose-response manner in cancer prevention by modulating telomeres through epigenetic mechanism.
Quiao et al [43]	Systematic review	16 (RCCTs) and 14 (CCTs)	—[g]	562	Myocardial infarction (1), healthy adults (4), obese (1), and polycystic ovary syndrome (1)	Combined training (1), endurance training (3), and aerobic exercise (3)	8 (1), 12 (1), 16 (2), 20 (1), 24 (1), and 48 (1)	qPCR	No	Yes	Weight-loss and comprehensive lifestyle intervention strategies show encouraging impacts in delaying telomere shortening. More rigorous studies targeting populations at different age stages through life span are needed.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Min et al [44]	Systematic review	2 (RCCTs)	3 (cross-sectional study)	247	2 (breast cancer survivor)	Aerobic exercise (2)	24 (1) and 48 (1)	qPCR	No	No	Three of the five studies reported that physical activity has a significant relationship in delaying TL shortening, but others observed no association between physical activity and TL in breast cancer survivors.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Marques et al [45]	Systematic review	4 (RCCTs)	16 (cross-sectional)	647	Obese (1) and healthy (3)	Aerobic exercise (4), HIIT (1), and endurance training (1)	24 (3) and 48 (1)	qPCR	No	Yes	Better cardiorespiratory fitness or a large cardiorespiratory training load are associated with an increase in TL. Although, TL was related to regular moderate-to-vigorous aerobic exercise and cardiorespiratory fitness in older healthy humans, it was not related to cardiorespiratory fitness among young subjects.

Study	Type of study	Study design, n		RCCTs[a] or CCTs[b] with physical exercise intervention					Risk of bias	Evaluation of quality	Conclusions
		Interventional	Observational	Sample		Intervention		Lab technique for TL[c]			
				n	Condition (number of studies, n)	Type (n)	Duration: weeks (n)				
Himbert et al [46]	Systematic review	10 (RCCTs) and 11 (CCTs)	—	439	Obese (1)	Aerobic exercise (1)	48 (1)	qPCR	No	No	The inconsistent effects of weight loss on telomere length or DNA repair suggest the need for a re-assessment of intervention designs and assay methodology to definitively address this topic.

[a]RCCT: randomized controlled clinical trial.

[b]CCT: (nonrandomized) controlled clinical trials.

[c]TL: telomere length.

[d]qPCR: polymerase chain reaction quantitative.

[e]HIIT: high-intensity interval training.

[f]NR: not reported.

[g]Not applicable.

AMSTAR-2 Appraisal

The overall confidence ranged from critically low to moderate quality scores. A total of 2 studies (16.6%) obtained an overall confidence of moderate, and 10 studies (83.4%) had critically low quality (Multimedia Appendix 2). The items with the highest scores were “did the review authors use a comprehensive literature search strategy?”, “did the review authors describe the included studies in adequate detail?”, “did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?”, and “did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?” The lowest scoring items were “did the review authors account for risk of bias in individual studies when interpreting/discussing the results of the review?”, “did the review authors report on the sources of funding for the studies included in the review?”, and “if meta-analysis was performed, did the review authors assess the potential impact of risk of bias in individual studies on the results of the meta-analysis or other evidence synthesis?” The interrater reliability of the methodological quality assessment was high (κ≥0.8).

ROBIS Assessment

Multimedia Appendix 3 summarizes the risk of bias using ROBIS. 40% (5/12) of the studies had a low risk of bias, 20% (2/12) had an unclear risk of bias, and 40% (5/12) had a low risk of bias. The domains related to “data collection” and “synthesis and findings” had the highest risk of bias. On the other hand, the domain related to “eligibility criteria” had the lowest risk of bias. The interrater reliability of the risk of bias was high (k>0.8).

PAGAC Grades of Evidence

The grades of evidence were classified as not assignable, limited, moderate, or strong according to the PAGAC. The level of evidence was limited to moderate among the studies. Most reviews evaluated a heterogeneous population of individuals, which limited the applicability and generalizability of the results. The domain “quantity and consistency” was affected because

some studies had inconsistency in the direction of or the effect size itself (Multimedia Appendix 4).

GROOVE Analysis

A total of 39 primary studies were identified across all systematic reviews and meta-analyses, 24 of which were distinct studies. The overall overlap in the matrix of evidence was moderate (CCA=32.5%) and it remained moderate (CCA=15.63%) after adjusting for the chronological structural missingness. Multimedia Appendix 5 shows a graphical representation of the GROOVE results.

Telomere Length

The number of studies in the systemic reviews that were finally integrated in the quantitative (meta-analytic analysis) was 22. The estimated effect size was 0.28 (95% CI 0.118-0.439), with a heterogeneity test value (Q) of 43.08 (P=.003). The I^2 coefficient was 51% and the prediction interval was −0.26 to 0.81. The number of weeks of intervention explained part of this heterogeneity (Q_B=7.54; P=.006). This factor explained 20% of the overall heterogeneity. Figure 2 shows that studies where the number of weeks of intervention was greater than 30 had smaller effect sizes. In fact, if the intervention time was categorized into <30 and ≥30 weeks, Q_B was 11.64 (P<.001).

Figure 2. Scatterplot of primary studies including effect size and number of weeks of intervention.

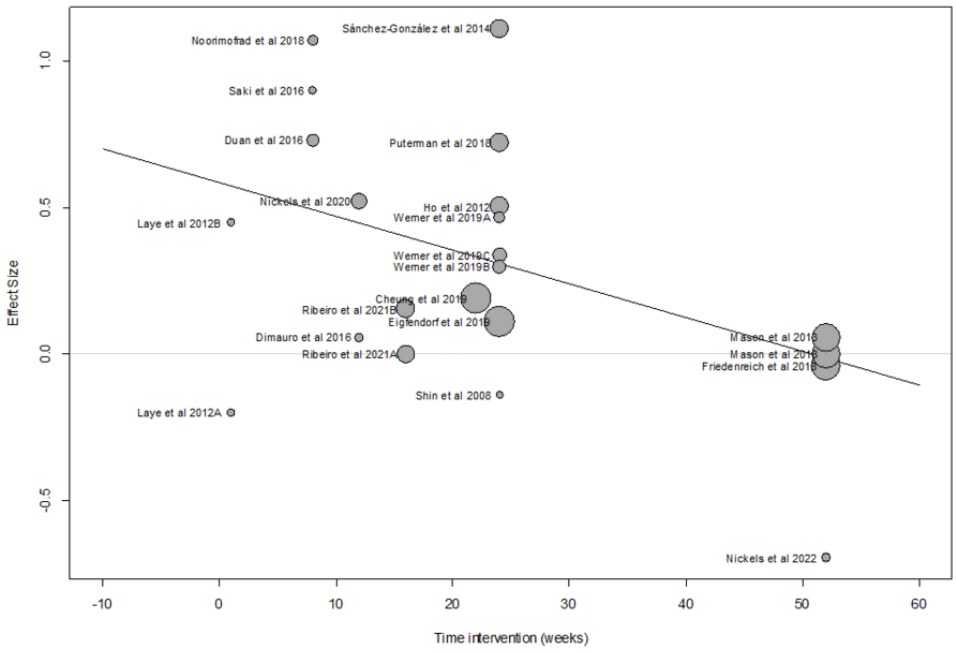
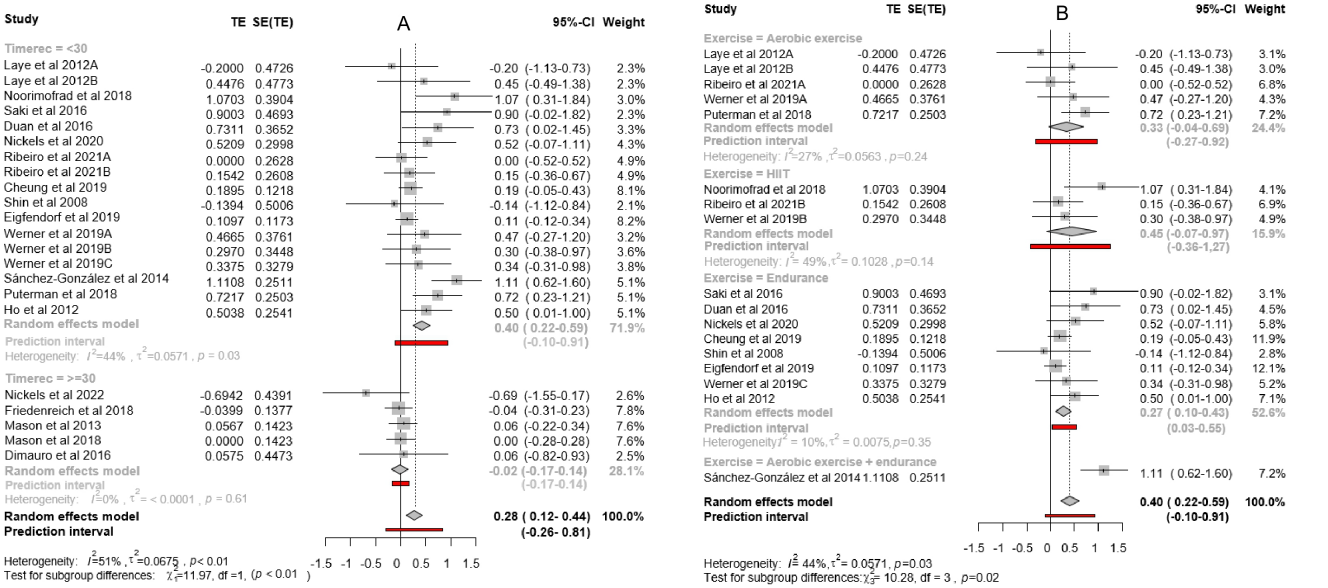


Figure 3A shows that the effect size was 0.39 for studies with an intervention length of less than 30 weeks, and −0.01 for those with interventions longer than 30 weeks. There was no significant difference in the average number of weekly sessions between these two groups (3.3 vs 3 sessions, respectively). Interestingly, the first group presented a statistically significant heterogeneity (I^2=41%), whilst the second, despite having a large range of outcomes, did not. Since heterogeneity was high in the group of studies with shorter interventions (<30 wk), it was analyzed with respect to other variables. The factor that better explained this heterogeneity was the type of physical exercise defined by the 4 categories (subgroups) as shown in Figure 3 (Q_B=11.15, P=.01). One subgroup was comprised of just 1 study that included a combination of aerobic exercise with endurance training and achieved a large effect size. The other 3 subgroups did not have a statistically significant heterogeneity between them, although the subgroup of studies defined by an intervention based on HIIT, which alternates short periods of intense anaerobic exercise with brief recovery periods, had a marked I^2 value (49%). The HIIT subgroup only included 3 studies, one of which had a very high effect size compared to the other two. When we compared the effect sizes of the 3 types of noncombined physical exercises, the value was small for aerobic exercise and endurance training, and moderate for HIIT.

Figure 3. (A) Forest plot of subgroups of primary studies defined by length of intervention. (B) Forest plot of subgroups of primary studies defined by type of exercise. TE: treatment effect.



The contour-enhanced funnel plot (Multimedia Appendix 6) did not identify publication bias. Many of the effect sizes were not significant, and at the top of the funnel, there was some asymmetry, which could be due to the heterogeneity of results. The nonpresence of publication bias was confirmed with the Egger method (*P*=.10). The Doi plot and LFK index also found a weak asymmetry (Multimedia Appendix 7).

Discussion

Principal Findings

This umbrella review studied the effects of physical exercise on TL. The estimated overall effect size indicated a small-moderate positive impact of exercise on TL. However, we found some discrepancies in the results across studies that may be explained by several methodological and contextual reasons, including different methodologies for measuring TL, such as quantitative polymerase chain reaction and terminal restriction fragment assays, that can contribute to inconsistencies in TL calculations. In this regard, Smith et al [47] pointed out that intraindividual TL can vary significantly depending on the lab technique used to measure it. Differences in study populations could also be a factor responsible for disparities in results. Demographic factors, such as age, sex, ethnicity, and health status, can influence the relationship between exercise and TL. For instance, LaRocca et al [48] and Puterman et al [21] suggest that the benefits of exercise on TL may be more pronounced in older adults and in those with higher stress levels. In addition, the duration and intensity of the intervention could be another confounding factor. Our findings suggest that interventions longer than 30 weeks tend to show a lesser positive impact on TL. This could be due to a ceiling effect where additional benefits of exercise do not translate into further increase of TL. This is consistent with Werner et al [29], who observed that the beneficial effects of exercise on TL are more evident in the early phases of physical exercise programs.

The high heterogeneity in study designs, including interventions, participant adherence, and control measures, could have also contributed to different outcomes. In fact, Denham et al [23] and Puterman et al [21] have already highlighted the importance of these methodological issues when interpreting TL results. Studies such as Cherkas et al [20] had found that vigorous aerobic physical activity was associated with longer telomeres in adults, which is in line with our finding of a positive effect of aerobic exercise, especially when combined with endurance training, on TL. However, a systematic review by Du et al [49] did not identify a significant association between aerobic exercise and TL in younger populations, suggesting that age may be an important moderating factor. Denham et al [23] found mixed results related to endurance physical training. Our review also found significant variation in the effects of noncombined endurance training, which suggests that factors such as the intensity and duration of this type of physical exercise may influence the magnitude of the effect on TL. Finally, HIIT had previously shown a significant positive effect on TL [29]. Our review results are consistent with this; nonetheless, the low number of HIIT studies and variety of designs limit generalizability.

Biological Understanding of the Effect of Exercise on TL

Reduction of Oxidative Stress and Inflammation

Oxidative stress and chronic inflammation are 2 major factors contributing to telomere erosion [50,51]. Studies have shown that regular exercise can mitigate these factors [52-54]. Acute and chronic physical activity have different effects on oxidative stress. While acute exercise triggers the production of reactive oxygen and nitrogen species, leading to oxidative stress, regular exercise training enhances the body's endogenous antioxidant system, offering protection against the harmful effects of oxidative damage [55]. This antioxidant effect protects cells from oxidative damage, which in turn preserves the integrity of telomeres. In addition, exercise reduces the levels of inflammatory markers such as C-reactive protein and tumor necrosis factor-alpha, which can protect telomeres from inflammation-induced damage [56].

Increase in Telomerase Activity

Telomerase is a crucial enzyme for maintaining TL, as it adds repetitive sequences to the end of chromosomes. Research suggests that exercise can increase telomerase activity. Ornish et al [57] demonstrated that an intensive lifestyle change program, which included regular exercise, significantly increased telomerase activity in men with prostate cancer. This finding suggests that exercise can directly influence telomere biology by activating telomerase, which helps maintain TL and protects cells from premature aging. In addition, a systematic review with meta-analysis published by Denham et al [23] concludes that regular exercise may attenuate telomere attrition through a telomerase-dependent mechanism and ultimately prolong lifespan and longevity. Furthermore, Ludlow et al [28] found that moderate physical activity levels are associated with increased telomerase activity in leukocytes, indicating that regular exercise not only improves overall health, but also acts at a molecular level to protect cells from premature aging. These studies highlight the importance of exercise continuity and intensity in maximizing the benefits related to telomerase activity and cellular longevity.

Improvement of Cardiovascular Capacity and Metabolic Health

Improved cardiovascular capacity and metabolic health are also associated with longer telomeres. Puterman et al [21] found that better aerobic capacity and higher insulin sensitivity are associated with longer telomeres. These cardiovascular and metabolic benefits reduce oxidative stress and inflammation, creating a healthier cellular environment that protects TL. Regular exercise enhances cardiovascular system efficiency, increases blood circulation, and facilitates the delivery of oxygen and nutrients to cells, which can contribute to the preservation of telomeres.

Strengths and Limitations

We present a comprehensive analysis of systematic reviews and meta-analyses of the effect of physical exercise on TL, providing a broad and robust synthesis of the available evidence. The methodology used to assess the quality of studies and the application of standardized tools for evaluating the risk of bias and heterogeneity strengthens the validity of our results. However, our work has some limitations. The methodological quality of the reviews included in this analysis varies, with many of them classified as critically low. Also, the heterogeneity in study designs, samples, and interventions makes difficult to generalize the results, even after the use of a random effects model to integrate outcomes. Furthermore, some meta-analyses reported results only using postintervention measures, whilst others used postintervention differences adjusted for preintervention measures; this may lead to variations in effect sizes.

Clinical Implications

The implementation of exercise programs into clinical practice could be an effective strategy to limit telomere shortening, promoting healthy aging and reducing the incidence of age-related diseases. HIIT and the combination of aerobic exercise with endurance training may be promising interventions in this regard but still require to be further tested in high-quality studies. In addition, physical exercise should be tailored to each individual's physical capability and health condition to maximize cellular health benefits. Implementing physical exercise programs will also require ongoing health education and monitoring to ensure adherence to and adjustment of interventions based on observed outcomes. Future studies should investigate sex and ethnic TL differences in response to various types of physical exercise. This could help personalize recommendations and optimize benefits on cellular health and aging.

Conclusions

Our umbrella review and meta-analysis identified a small-moderate positive effect of physical exercise on TL, which seems to be influenced by the duration and type of physical exercise. In health care systems, the implementation of evidence-based physical exercise, training programs, and/or recommendations tailored to each individual might be a valuable preventive strategy for a healthy aging. More studies, with larger sample sizes, testing the impact of standardized, evidence-based physical exercise interventions on TL are warranted.

Disclaimer

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. Dissemination to participants and related patients and public communities: After publication, the findings of this review will be disseminated to appropriate audiences, such as academia, clinicians, policy makers, and the general public, through various channels including blogs, press releases, and social media.

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JLS-G, JLS-R, RG-S, JP, and JM-V contributed to the design of this review. JLS-G, VN-L, JP, and JM-V conducted study selection, data extraction and analysis, and wrote the first draft of this manuscript. RJ-V provided a detailed review of the statistical

analysis. All authors reviewed and approved the final version of the manuscript. The corresponding author, JP, confirms that all listed authors meet authorship criteria and that no others meeting these criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Database formulas during literature search.
[DOCX File, 13 KB - aging_v8i1e64539_app1.docx]

Multimedia Appendix 2
Risk of bias.
[PNG File, 263 KB - aging_v8i1e64539_app2.png]

Multimedia Appendix 3
Graphical representation of the overlap of primary studies between reviews.
[PNG File, 100 KB - aging_v8i1e64539_app3.png]

Multimedia Appendix 4
Contour-enhanced funnel plot of the primary studies included in the meta-analysis.
[PNG File, 22 KB - aging_v8i1e64539_app4.png]

Multimedia Appendix 5
Doi plot and LFK index of the primary studies included in the meta-analysis.
[PNG File, 7 KB - aging_v8i1e64539_app5.png]

Multimedia Appendix 6
Quality assessment scores (A Measurement Tool to Assess Systematic Reviews [AMSTAR-2]).
[DOCX File, 18 KB - aging_v8i1e64539_app6.docx]

Multimedia Appendix 7
Summary of findings and quality of evidence (Physical Activity Guidelines Advisory Committee Grading Criteria [PAGAC]).
[DOCX File, 16 KB - aging_v8i1e64539_app7.docx]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File, 83 KB - aging_v8i1e64539_app8.pdf]

References

1. World Population Prospects 2022: Summary of results. : United Nations; 2022.
2. Melzer D, Pilling LC, Ferrucci L. The genetics of human ageing. Nat Rev Genet 2020 Feb;21(2):88-101. [doi: 10.1038/s41576-019-0183-6] [Medline: 31690828]
3. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. Cell 2023 Jan 19;186(2):243-278. [doi: 10.1016/j.cell.2022.11.001] [Medline: 36599349]
4. Zakian VA. Telomeres: beginning to understand the end. Science 1995 Dec 8;270(5242):1601-1607. [doi: 10.1126/science.270.5242.1601] [Medline: 7502069]
5. Smith EM, Pendlebury DF, Nandakumar J. Structural biology of telomeres and telomerase. Cell Mol Life Sci 2020 Jan;77(1):61-79. [doi: 10.1007/s00018-019-03369-x] [Medline: 31728577]
6. Opresko PL, Shay JW. Telomere-associated aging disorders. Ageing Res Rev 2017 Jan;33:52-66. [doi: 10.1016/j.arr.2016.05.009] [Medline: 27215853]
7. Cascales-Angosto M, Álvarez-Gómez JA. Telómeros y telomerasa, sus implicaciones en el envejecimiento y el cáncer [Article in Spanish]. An Real Acad Dr Esp 2010;14:49-70.
8. Erdem HB, Bahsi T, Ergün MA. Function of telomere in aging and age related diseases. Environ Toxicol Pharmacol 2021 Jul;85:103641. [doi: 10.1016/j.etap.2021.103641] [Medline: 33774188]
9. Armanios M. The role of telomeres in human disease. Annu Rev Genomics Hum Genet 2022 Aug 31;23:363-381. [doi: 10.1146/annurev-genom-010422-091101] [Medline: 35609925]

10. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013 Apr;45(4):422-427. [doi: 10.1038/ng.2528] [Medline: 23535734]
11. Mangino M, Hwang SJ, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet 2012 Dec 15;21(24):5385-5394. [doi: 10.1093/hmg/dds382] [Medline: 23001564]
12. Willeit P, Willeit J, Mayr A, et al. Telomere length and risk of incident cancer and cancer mortality. JAMA 2010 Jul 7;304(1):69-75. [doi: 10.1001/jama.2010.897] [Medline: 20606151]
13. Diez Roux AV, Ranjit N, Jenny NS, et al. Race/ethnicity and telomere length in the Multi-Ethnic Study of Atherosclerosis. Aging Cell 2009 Jun;8(3):251-257. [doi: 10.1111/j.1474-9726.2009.00470.x] [Medline: 19302371]
14. Epel ES. Psychological and metabolic stress: a recipe for accelerated cellular aging? HORMONES (Athens) 2009;8(1):7-22. [doi: 10.14310/horm.2002.1217] [Medline: 19269917]
15. Simon NM, Smoller JW, McNamara KL, et al. Telomere shortening and mood disorders: preliminary support for a chronic stress model of accelerated aging. Biol Psychiatry 2006 Sep 1;60(5):432-435. [doi: 10.1016/j.biopsych.2006.02.004] [Medline: 16581033]
16. Valdes A, Andrew T, Gardner J, et al. Obesity, cigarette smoking, and telomere length in women. Lancet 2005 Aug;366:662-664. [doi: 10.1016/S0140-6736(05)66630-5]
17. Pavanello S, Hoxha M, Dioni L, et al. Shortened telomeres in individuals with abuse in alcohol consumption. Int J Cancer 2011 Aug 15;129(4):983-992. [doi: 10.1002/ijc.25999] [Medline: 21351086]
18. McGrath M, Wong JYY, Michaud D, Hunter DJ, De Vivo I. Telomere length, cigarette smoking, and bladder cancer risk in men and women. Cancer Epidemiol Biomarkers Prev 2007 Apr;16(4):815-819. [doi: 10.1158/1055-9965.EPI-06-0961] [Medline: 17416776]
19. Sánchez-González JL, Sánchez-Rodríguez JL, Martín-Vallejo J, Martel-Martel A, González-Sarmiento R. Effects of physical exercise on cognition and telomere length in healthy older women. Brain Sci 2021 Oct 27;11(11):1417. [doi: 10.3390/brainsci11111417] [Medline: 34827416]
20. Cherkas LF, Hunkin JL, Kato BS, et al. The association between physical activity in leisure time and leukocyte telomere length. Arch Intern Med 2008 Jan 28;168(2):154-158. [doi: 10.1001/archinternmed.2007.39] [Medline: 18227361]
21. Puterman E, Lin J, Blackburn E, O'Donovan A, Adler N, Epel E. The power of exercise: buffering the effect of chronic stress on telomere length. PLoS One 2010 May 26;5(5):e10837. [doi: 10.1371/journal.pone.0010837] [Medline: 20520771]
22. Buttet M, Bagheri R, Ugbolue UC, et al. Effect of A lifestyle intervention on telomere length: a systematic review and meta-analysis. Mech Ageing Dev 2022 Sep;206:111694. [doi: 10.1016/j.mad.2022.111694] [Medline: 35760212]
23. Denham J, Sellami M. Exercise training increases telomerase reverse transcriptase gene expression and telomerase activity: a systematic review and meta-analysis. Ageing Res Rev 2021 Sep;70:101411. [doi: 10.1016/j.arr.2021.101411] [Medline: 34284150]
24. Mundstock E, Zatti H, Louzada FM, et al. Effects of physical activity in telomere length: systematic review and meta-analysis. Ageing Res Rev 2015 Jul;22:72-80. [doi: 10.1016/j.arr.2015.02.004] [Medline: 25956165]
25. Valente C, Andrade R, Alvarez L, Rebelo-Marques A, Stamatakis E, Espregueira-Mendes J. Effect of physical activity and exercise on telomere length: systematic review with meta-analysis. J Am Geriatr Soc 2021 Nov;69(11):3285-3300. [doi: 10.1111/jgs.17334] [Medline: 34161613]
26. Song S, Lee E, Kim H. Does exercise affect telomerelength? a systematic review and meta-analysis of randomized controlled trials. Medicina (Kaunas) 2022 Feb 5;58(2):242. [doi: 10.3390/medicina58020242] [Medline: 35208566]
27. Denham J, O'Brien BJ, Charchar FJ. Telomere length maintenance and cardio-metabolic disease prevention through exercise training. Sports Med 2016 Sep;46(9):1213-1237. [doi: 10.1007/s40279-016-0482-4] [Medline: 26914269]
28. Ludlow AT, Zimmerman JB, Witkowski S, Hearn JW, Hatfield BD, Roth SM. Relationship between physical activity level, telomere length, and telomerase activity. Med Sci Sports Exerc 2008;40(10):1764-1771. [doi: 10.1249/MSS.0b013e31817c92aa]
29. Werner CM, Hecksteden A, Morsch A, et al. Differential effects of endurance, interval, and resistance training on telomerase activity and telomere length in a randomized, controlled study. Eur Heart J 2019 Jan 1;40(1):34-46. [doi: 10.1093/eurheartj/ehy585] [Medline: 30496493]
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
31. Stone PW. Popping the (PICO) question in research and evidence-based practice. Appl Nurs Res 2002 Aug;15(3):197-198. [doi: 10.1053/apnr.2002.34181] [Medline: 12173172]
32. Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998 Aug;352:609-613. [doi: 10.1016/S0140-6736(98)01085-X]
33. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017 Sep 21;358:j4008. [doi: 10.1136/bmj.j4008] [Medline: 28935701]
34. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]

35. Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016 Jan;69:225-234. [doi: 10.1016/j.jclinepi.2015.06.005] [Medline: 26092286]
36. Torres A, Tennant B, Ribeiro-Lucas I, Vaux-Bjerke A, Piercy K, Bloodgood B. Umbrella and systematic review methodology to support the 2018 Physical Activity Guidelines Advisory Committee. J Phys Act Health 2018 Nov 1;15(11):805-810. [doi: 10.1123/jpah.2018-0372] [Medline: 30336718]
37. Bracchiglione J, Meza N, Bangdiwala S, et al. Graphical Representation of Overlap for OVErviews. : OSF; 2022. [doi: 10.17605/OSF.IO/U2MS4]
38. Skipka G. The inclusion of the estimated inter-study variation into forest plots for random effects meta-analysis - a suggestion for a graphical representation. In: Program and Abstract Book: Cochrane Colloquium; 2006:23-26.
39. Ortega Z, Martín-Vallejo J, Mencía A, Galindo-Villardón MP, Pérez-Mellado V. Introducing meta-partition, a useful methodology to explore factors that influence ecological effect sizes. PLoS One 2016;11(7):e0158624. [doi: 10.1371/journal.pone.0158624] [Medline: 27409084]
40. Barragán R, Ortega-Azorín C, Sorlí JV, et al. Effect of physical activity, smoking, and sleep on telomere length: a systematic review of observational and intervention studies. J Clin Med 2021 Dec 24;11(1):76. [doi: 10.3390/jcm11010076] [Medline: 35011817]
41. Schellnegger M, Lin AC, Hammer N, Kamolz LP. Physical activity on telomere length as a biomarker for aging: a systematic review. Sports Med Open 2022 Sep 4;8(1):111. [doi: 10.1186/s40798-022-00503-1] [Medline: 36057868]
42. Prathap L, Suganthirababu SP, Kumar KP, S P. Effect of physical exercise in remodeling telomere length and cancer prevention in an epigenetic prospect – a systematic review. Biomed Pharmacol J 2021 Jun 30;14(2):891-901. [doi: 10.13005/bpj/2191]
43. Qiao S, Jiang Y, Li X. The impact of health promotion interventions on telomere length: a systematic review. Am J Health Promot 2020 Jul;34(6):633-647. [doi: 10.1177/0890117120906958] [Medline: 32103672]
44. Min J, Kim JY, Choi JY, Kong ID. Association between physical activity and telomere length in women with breast cancer: a systematic review. J Clin Med 2022 Apr 30;11(9):2527. [doi: 10.3390/jcm11092527] [Medline: 35566652]
45. Marques A, Gouveira É, Peralta M, et al. Cardiorespiratory fitness and telomere length: a systematic review. J Sports Sci 2020 Jul;38(14):1690-1697. [doi: 10.1080/02640414.2020.1754739] [Medline: 32284029]
46. Himbert C, Thompson H, Ulrich CM. Effects of intentional weight loss on markers of oxidative stress, DNA repair and telomere length - a systematic review. Obes Facts 2017;10(6):648-665. [doi: 10.1159/000479972] [Medline: 29237161]
47. Smith DL, Wu C, Gregorich S, Dai G, Lin J. Impact of DNA extraction methods on quantitative PCR telomere length assay precision in human saliva samples. Int J Methodol 2022;1(1):44-57. [doi: 10.21467/ijm.1.1.5784]
48. LaRocca TJ, Seals DR, Pierce GL. Leukocyte telomere length is preserved with aging in endurance exercise-trained adults and related to maximal aerobic capacity. Mech Ageing Dev 2010 Feb;131(2):165-167. [doi: 10.1016/j.mad.2009.12.009] [Medline: 20064545]
49. Du M, Prescott J, Kraft P, et al. Physical activity, sedentary behavior, and leukocyte telomere length in women. Am J Epidemiol 2012 Mar 1;175(5):414-422. [doi: 10.1093/aje/kwr330] [Medline: 22302075]
50. Finkel T, Holbrook NJ. Oxidants, oxidative stress and the biology of ageing. Nat New Biol 2000 Nov 9;408(6809):239-247. [doi: 10.1038/35041687] [Medline: 11089981]
51. von Zglinicki T. Oxidative stress shortens telomeres. Trends Biochem Sci 2002 Jul;27(7):339-344. [doi: 10.1016/S0968-0004(02)02110-2]
52. Powers SK, Jackson MJ. Exercise-induced oxidative stress: cellular mechanisms and impact on muscle force production. Physiol Rev 2008 Oct;88(4):1243-1276. [doi: 10.1152/physrev.00031.2007] [Medline: 18923182]
53. Gomez-Cabrera MC, Domenech E, Viña J. Moderate exercise is an antioxidant: upregulation of antioxidant genes by training. Free Radic Biol Med 2008 Jan 15;44(2):126-131. [doi: 10.1016/j.freeradbiomed.2007.02.001] [Medline: 18191748]
54. Radak Z, Chung HY, Koltai E, Taylor AW, Goto S. Exercise, oxidative stress and hormesis. Ageing Res Rev 2008 Jan;7(1):34-42. [doi: 10.1016/j.arr.2007.04.004] [Medline: 17869589]
55. Belviranlı M, Gökbel H. Acute exercise induced oxidative stress and antioxidant changes. Electron J Gen Med 2006;3(3):126-131. [doi: 10.29333/ejgm/82392]
56. Fischer CP. Interleukin-6 in acute exercise and training: what is the biological relevance? Exerc Immunol Rev 2006;12:6-33. [Medline: 17201070]
57. Ornish D, Lin J, Daubenmier J, et al. Increased telomerase activity and comprehensive lifestyle changes: a pilot study. Lancet Oncol 2008 Nov;9(11):1048-1057. [doi: 10.1016/S1470-2045(08)70234-1] [Medline: 18799354]

Abbreviations

AMSTAR-2: A Measurement Tool to Assess Systematic Reviews
CCA: corrected covered area
GROOVE: Graphical Representation of Overlap for OVErviews
HIIT: high-intensity interval training
PAGAC: Physical Activity Guidelines Advisory Committee Grading Criteria

PICOS: Population, Intervention, Control, Outcomes, and Study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ROBIS: Risk Of Bias in Systematic reviews
TL: telomere length

Edited by M O'Connell; submitted 19.07.24; peer-reviewed by MN Shalaby, R Sobral-Monteiro-Junior; revised version received 10.09.24; accepted 16.10.24; published 10.01.25.

Please cite as:
Sánchez-González JL, Sánchez-Rodríguez JL, González-Sarmiento R, Navarro-López V, Juárez-Vela R, Pérez J, Martín-Vallejo J
Effect of Physical Exercise on Telomere Length: Umbrella Review and Meta-Analysis
JMIR Aging 2025;8:e64539
URL: https://aging.jmir.org/2025/1/e64539
doi:10.2196/64539

Original Paper

Exploring the Feasibility of a 5-Week mHealth Intervention to Enhance Physical Activity and an Active, Healthy Lifestyle in Community-Dwelling Older Adults: Mixed Methods Study

Kim Daniels[1,2], MSc; Sharona Vonck[1], PhD; Jolien Robijns[1], PhD; Kirsten Quadflieg[1], PhD; Jochen Bergs[1,3], Prof Dr; Annemie Spooren[1,2], Prof Dr; Dominique Hansen[2,4], Prof Dr; Bruno Bonnechère[1,2,5], Prof Dr

[1]Centre of Expertise in Care Innovation, Department of PXL – Healthcare, PXL University of Applied Sciences and Arts, Hasselt, Belgium

[2]REVAL Rehabilitation Research Center, Faculty of Rehabilitation Sciences, Hasselt University, Diepenbeek, Belgium

[3]THINK3 Simulation & Innovation Lab, Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium

[4]BIOMED, Faculty of Medicine and Life Sciences, Hasselt University, Diepenbeek, Belgium

[5]Technology-Supported and Data-Driven Rehabilitation, Data Sciences Institute, Hasselt University, Diepenbeek, Belgium

Corresponding Author:
Kim Daniels, MSc
Centre of Expertise in Care Innovation, Department of PXL – Healthcare
PXL University of Applied Sciences and Arts
Guffenslaan 39
Hasselt, 3500
Belgium
Phone: 32 485763451
Email: kim.daniels@pxl.be

Abstract

Background: Advancements in mobile technology have paved the way for innovative interventions aimed at promoting physical activity (PA).

Objective: The main objective of this feasibility study was to assess the feasibility, usability, and acceptability of the More In Action (MIA) app, designed to promote PA among older adults. MIA offers 7 features: personalized tips, PA literacy, guided peer workouts, a community calendar, a personal activity diary, a progression monitor, and a chatbot.

Methods: Our study used a mixed methods approach to evaluate the MIA app's acceptability, feasibility, and usability. First, a *think-aloud* method was used to provide immediate feedback during initial app use. Participants then integrated the app into their daily activities for 5 weeks. Behavioral patterns such as user session duration, feature use frequency, and navigation paths were analyzed, focusing on engagement metrics and user interactions. User satisfaction was assessed using the System Usability Scale, Net Promoter Score, and Customer Satisfaction Score. Qualitative data from focus groups conducted after the 5-week intervention helped gather insights into user experiences. Participants were recruited using a combination of web-based and offline strategies, including social media outreach, newspaper advertisements, and presentations at older adult organizations and local community services. Our target group consisted of native Dutch-speaking older adults aged >65 years who were not affected by severe illnesses. Initial assessments and focus groups were conducted in person, whereas the intervention itself was web based.

Results: The study involved 30 participants with an average age of 70.3 (SD 4.8) years, of whom 57% (17/30) were female. The app received positive ratings, with a System Usability Scale score of 77.4 and a Customer Satisfaction Score of 86.6%. Analysis showed general satisfaction with the app's workout videos, which were used in 585 sessions with a median duration of 14 (IQR 0-34) minutes per day. The Net Promoter Score was 33.34, indicating a good level of customer loyalty. Qualitative feedback highlighted the need for improvements in navigation, content relevance, and social engagement features, with suggestions for better calendar visibility, workout customization, and enhanced social features. Overall, the app demonstrated high usability and satisfaction, with near-daily engagement from participants.

Conclusions: The MIA app shows significant potential for promoting PA among older adults, evidenced by its high usability and satisfaction scores. Participants engaged with the app nearly daily, particularly appreciating the workout videos and educational content. Future enhancements should focus on better calendar visibility, workout customization, and integrating social networking features to foster community and support. In addition, incorporating wearable device integration and predictive analytics could

provide real-time health data, optimizing activity recommendations and health monitoring. These enhancements will ensure that the app remains user-friendly, relevant, and sustainable, promoting sustained PA and healthy behaviors among older adults.

Trial Registration: ClinicalTrials.gov NCT05650515; https://clinicaltrials.gov/study/NCT05650515

(JMIR Aging 2025;8:e63348) doi:10.2196/63348

KEYWORDS

mobile health; mHealth; feasibility; physical activity; older adults; health promotion; usability; mobile phone

Introduction

Background

Ensuring the well-being of the global aging population is a pressing concern, particularly as the World Health Organization predicts a significant rise in the number of individuals aged >65 years by 2050 [1]. To confront this impending public health challenge, the promotion of regular physical activity (PA) is paramount [2].

Research has consistently demonstrated the multifaceted advantages of regular PA on physical, cognitive, and mental health even in advanced age. In recognition of these benefits, the World Health Organization advocates for specific PA guidelines for individuals aged ≥65 years [3,4]. These guidelines encompass moderate-intensity aerobic activities, muscle-strengthening exercises, and balance training. However, despite the wealth of evidence supporting these recommendations, a substantial proportion of older adults worldwide fail to meet the prescribed PA levels [5,6]. Hence, adherence to recommended guidelines remains a challenging issue [7].

Addressing the gap in PA participation among older adults necessitates the implementation of innovative and sustainable interventions. Mobile technology has opened up avenues for innovative approaches to foster PA and cultivate healthier lifestyles even among older adults [8,9]. Mobile health (mHealth) apps have emerged as promising and cost-effective health intervention tools, especially in promoting PA. These apps harness technology to offer personalized guidance and track and encourage PA [10]. Capitalizing on the widespread use of smartphones and tablets, these apps possess the potential to revolutionize how we address and promote PA among aging populations [11]. They offer convenient access to PA advice and support, deliver an enjoyable user experience, and furnish feedback on progress over time [12-14].

Despite these positive trends, adoption of these technologies in real-life conditions is still relatively limited. Several challenges and hurdles must be overcome before such interventions can be successfully and sustainably implemented in the field. The challenge is not solely in developing mHealth apps but also in maintaining long-term engagement and motivation among older adults. In addition, studies often report high attrition rates and small intervention effects [14-16]. These problems must be addressed to use the full potential of mHealth interventions. First, high attrition is likely due to the intervention not matching the users' needs, goals, and expectations [17,18]. This may be avoided by involving potential users during the entire cycle of intervention development [19-22]. Second, research has revealed that theory-based interventions are more effective at modifying health behaviors than traditional interventions [23,24]. Thus, interventions should be grounded within and informed by theoretical models.

Objectives

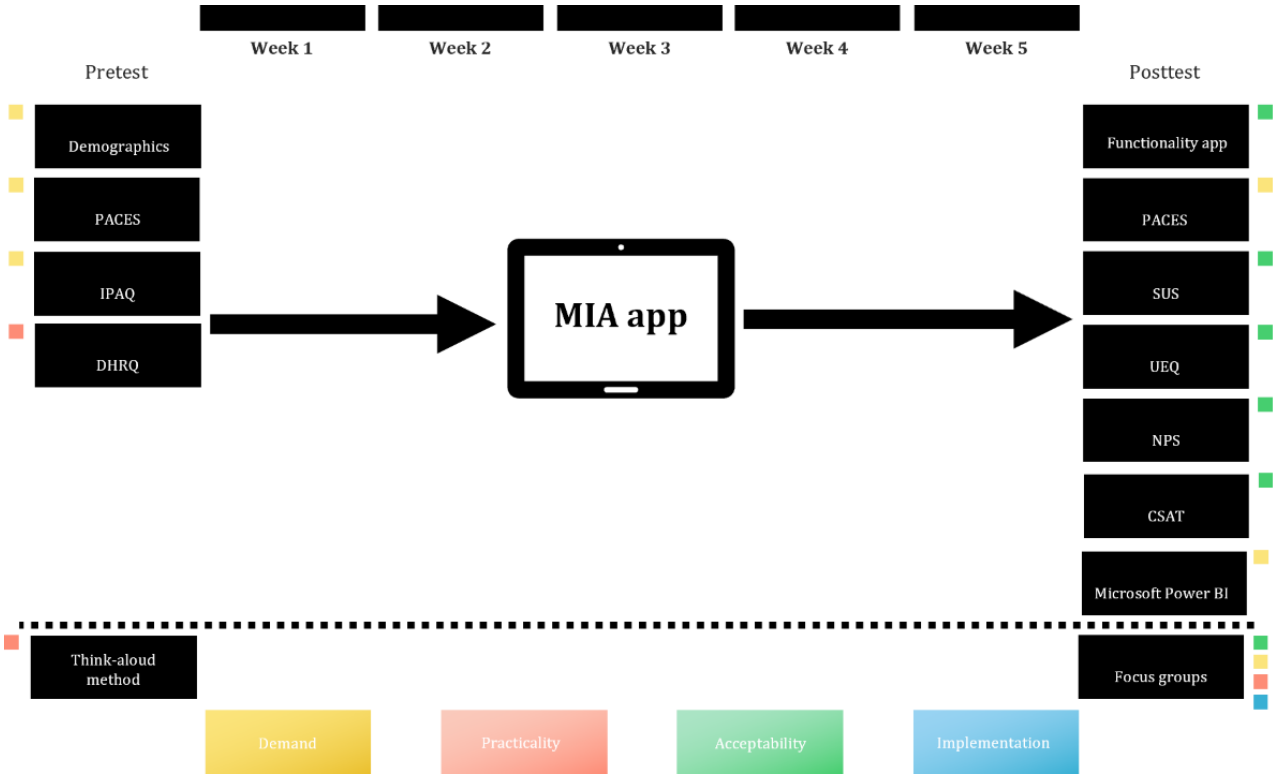
In response to these challenges, a collaborative cocreation process led to the development and refinement of an mHealth app named More In Action (MIA) [19]. The app's content and design were crafted through cocreative workshops and based on the theoretical framework of the Behavior Change Wheel (BCW) [23]. Despite the active involvement of end users from the outset of development and throughout the iterative process, as well as preliminary results indicating high levels of enjoyment and ease of use of the app [19], the long-term retention rate and motivation levels of the participants remain unknown. Therefore, the primary objective of this study was to thoroughly explore the acceptability, feasibility, and usability of the MIA app in promoting PA and encouraging sustained, active, and healthy behaviors among its users.

Methods

Study Design

We used a comprehensive mixed methods approach to assess the MIA app's acceptability, feasibility, and usability, as shown in Figure 1.

Figure 1. Study design. The colors represent the different evaluation focuses: yellow for demand (interest and need), orange for practicality (feasibility), green for acceptability (user satisfaction and acceptance), and blue for implementation (integration into the study). CSAT: Customer Satisfaction Score; DHRQ: Digital Health Readiness Questionnaire; IPAQ: International Physical Activity Questionnaire; MIA: More In Action; NPS: Net Promoter Score; PACES: Physical Activity Enjoyment Scale; SUS: System Usability Scale; UEQ: User Experience Questionnaire.



Our methodology adhered to the recommended framework for feasibility studies by Bowen et al [25]. This framework allowed us to explore critical aspects of the MIA app, such as *acceptability*, *demand*, *practicality*, and *implementation*. These interrelated concepts were systematically integrated into our study design and procedures, as detailed in Table 1.

Acceptability was defined as the perceived suitability of the app and is directly tied to user satisfaction and their intent to continue using the app, as shown in Table 1. Feasibility, on the other hand, addresses the ease of implementing the app among older adults [26], focusing on demand and implementation and examining whether it can be effectively implemented and sustained in users' daily routines. In Table 1, feasibility is broken down into 2 distinct areas. The first is demand, which addresses whether older adults are willing to adopt and regularly use the app. This area evaluates actual use, intention to use, and perceived demand. The second is practicality, which assesses how easily the MIA app can be integrated into the users' daily lives. This area was explored through the *think-aloud* method, focus groups, and questionnaires. Practicality measures the ease of integrating the app into daily routines, user willingness to pay, and any perceived positive or negative effects. Usability, a key aspect of our investigation, was defined as the ease of use and suitability of the system or product for a specific user group performing designated tasks in a particular environment, including practicality. In this context, *ease of use* directly impacted user performance and satisfaction, whereas *acceptability* determined the likelihood of the product being embraced and used [27]. In Table 1, usability directly impacts acceptability (satisfaction and intent to continue using the app) and feasibility (ease of integration and ability to perform the required intervention activities).

Table 1. Key areas of focus for the More In Action (MIA) feasibility study based on the framework by Bowen et al [25].

Area of focus and research question	Method	Outcome	Measures
Usability			
Acceptability			
To what extent is the MIA app as a means to promote PA[a] acceptable among community-dwelling adults aged ≥65 years?	• Questionnaire • Focus group	• Satisfaction • Intent to continue use	SUS[b], CSAT[c], NPS[d], and Likert-scale questions on how the app fits into the end users' daily-life activities
Feasibility			
Demand			
What is the level of adoption of the MIA app among community-dwelling older adults? What factors influence the intention to use and engage with the app?	• Questionnaire • Microsoft Power BI (Microsoft Corp) analytics • Plausible Analytics	• Actual use • Intention to use • Perceived demand	Questions regarding the influence of the app on modulating PA behavior (eg, PACES[e] and IPAQ[f]) and analytics to compare the frequency of use and patterns of use across the participants
Practicality			
To what extent can the MIA app be integrated into the daily lives of older adults aged ≥65 years residing within the community?	• Think-aloud method • Questionnaire • Focus group	• Positive or negative effects • Ability of participants to execute intervention activities • Willingness to pay	First impressions, Likert scale, and in-depth questions regarding integrating the app into their daily lives (eg, DHRQ[g])
Implementation			
How can the MIA app be optimally implemented to facilitate sustained engagement in PA among community-dwelling older adults?	• Focus group	• Amount and type of resources needed to implement • Factors affecting implementation	In-depth questions on how the MIA app can be deployed in the community context

[a]PA: physical activity.

[b]SUS: System Usability Scale.

[c]CSAT: Customer Satisfaction Score.

[d]NPS: Net Promoter Score.

[e]PACES: Physical Activity Enjoyment Scale.

[f]IPAQ: International Physical Activity Questionnaire.

[g]DHRQ: Digital Health Readiness Questionnaire.

Ethical Considerations

This study was registered at Clinical Trials.gov (NCT05650515) and was approved by the ethical committee of Hasselt University (B1152023000011). All participants provided informed consent. To ensure privacy, the data used in this study were deidentified before analysis. Participants did not receive any compensation for their involvement.

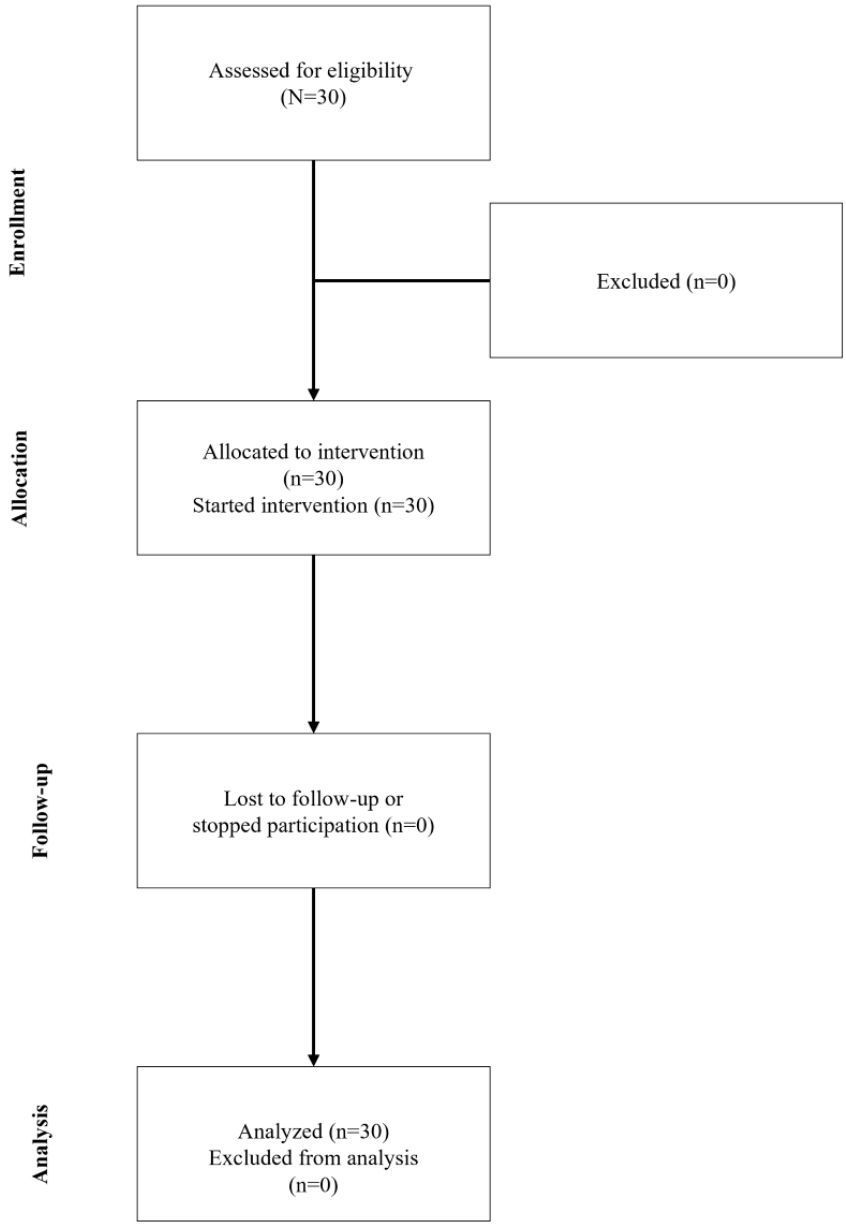
Participants

Older adults without severe illness were invited to participate in this study. They were recruited via social media outreach, newspaper advertisements, and pitches at several older adult organizations and through the local community services during the recruitment period from August 2023 to September 2023.

The inclusion criteria for participants encompassed that they had to be aged ≥65 years, competent to provide informed consent (Multimedia Appendix 1), able to actively participate in the study, community dwelling (living either independently at home or in a serviced apartment), without any severe illnesses, and native Dutch speakers. The exclusion criteria were the presence of current neurological, cardiovascular, respiratory, severe metabolic, or cognitive disorders. Participants were not excluded based on digital literacy, ensuring a diverse range of digital competence levels. The complete list of exclusion criteria is presented in Multimedia Appendix 2.

Figure 2 provides a CONSORT (Consolidated Standards of Reporting Trials) flowchart for this single-arm feasibility study. The methodology and results were reported following the CONSORT 2010 checklist [28].

Figure 2. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram for a single-arm feasibility study of the 5-week More In Action intervention.



Intervention

General Description

The mHealth app MIA represents an innovative approach to promoting PA and fostering a lifestyle centered on health and activity among older adults. Developed through a collaborative cocreation process, MIA has been refined over several iterations to meet the specific needs and preferences of its target user group [19]. MIA is optimized for both smartphone and tablet use, thereby ensuring accessibility and user engagement across a broad spectrum of mobile devices.

The app's design and content draw upon insights gained from cocreative workshops and are firmly grounded in the intervention functions of the BCW theoretical framework [29] as well as principles from self-identification theory [30,31] as key elements to motivating individuals toward healthier behaviors. Textbox 1 presents a comprehensive summary of the behavior change techniques incorporated into MIA.

Textbox 1. Overview of behavior change techniques implemented in the More In Action (MIA) app.

Behavior Change Wheel (BCW) [32]

- The BCW framework is centered on 3 core components: capability, opportunity, and motivation (Capability, Opportunity, and Motivation–Behavior system), which can drive behavior change. The MIA app enhances users' *capability* through educational content and reminders, increases *opportunity* through social features such as the community calendar, and boosts *motivation* by offering personalized goals and feedback.

Self-determination theory (SDT) [33]

- SDT emphasizes 3 fundamental psychological needs—autonomy, competence, and relatedness—as essential for motivation toward healthy behavior changes:
- *Autonomy*: the MIA app can provide users with the ability to customize their exercise routines, choose their goals, and select the types of physical activity (PA) they prefer. The workouts are composed personally based on the tailored personal goals users enter after registering.
- *Competence*: to enhance users' feelings of competence, the MIA app incorporates a system of progressive challenges and feedback. The app tracks users' progress in real time, offering progress bars or celebratory messages for achieving milestones (eg, achieving the World Health Organization guideline for PA). In addition, educational content on the benefits of regular PA and other health benefits helps users feel more skilled and capable.
- *Relatedness*: the MIA app fosters a sense of relatedness by integrating a community calendar that allows users to connect with others.

Self-identification theory (SIT) [31]

- Implementing SIT within the MIA app involved creating features that allow older adults to integrate PA into their self-concept, making it a core part of their identity. Users start by *selecting goals* that resonate with their personal aspirations and lifestyle. These goals can range from improving health, gaining strength, and enhancing mobility to participating in community activities or playing with grandchildren. They choose freely. The key is for users to choose goals that reflect their values and how they perceive themselves.

Unique Attributes of MIA

The MIA app exhibits several unique attributes within the scope of gerontechnology that set it apart from other existing mHealth technologies. To compare MIA with existing solutions, we conducted a review to map the currently available mHealth technologies for promoting PA in older adults.

All technologies were scored according to 11 key components previously identified [19,20,34]. These findings are detailed in Table 2.

Table 2. Comparative analysis of mobile health app features for enhancing physical activity among older adults.

Existing apps	User-centered design[a]	Behavior change techniques[b]	Personalized intervention[c]	Interactivity[d]	Activities of daily living[e]	Integration with wearable devices[f]	Social cohesion[g]	Education and information[h]	Rewards and incentives[i]	Accessibility and inclusion[j]	Older adult–specific features[k]
Standing-Tall [35]			✓[l]	✓						✓	✓
Web + [36]		✓	✓			✓	✓	✓	✓	✓	✓
Ready Steady Go [37]		✓		✓			✓	✓	✓	✓	✓
HBex [24]	✓	✓	✓	✓	✓			✓	✓	✓	
Vivo [38]	✓			✓						✓	
App-based exercise program [39]			✓		✓			✓		✓	✓
Physitrack [40]			✓	✓						✓	
My plan 2.0 [41]	✓	✓	✓	✓				✓	✓	✓	✓
Bingocize [42]				✓			✓	✓	✓	✓	✓
Fit for All platform [43]			✓	✓					✓		
ActiveLifestyle [44]		✓	✓					✓		✓	✓
Vibrotactile app [45]	✓		✓	✓					✓		✓
PACE[m] app [46]			✓		✓	✓			✓	✓	
Gymcentral [47]	✓	✓	✓	✓			✓		✓	✓	✓
Telehealth intervention [48]			✓	✓			✓				
Nymbl [49]			✓	✓				✓	✓		✓
eLIFE [50]		✓	✓	✓				✓	✓	✓	✓
Exercise app [51]			✓	✓					✓	✓	✓
Make Movement Your Mission [52]		✓		✓	✓		✓	✓	✓	✓	✓
Evident [53]		✓	✓	✓		✓		✓	✓		✓
MIA[n] app [19]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[a]This design prioritizes the end user's needs, preferences, and limitations throughout the development process.

[b]These are systematic strategies derived from behavioral science theories to influence and sustain behavior modification.

[c]Personalized interventions involve tailoring health-related strategies and communications to individual users based on specific data gathered about their behaviors, preferences, and environmental contexts.

[d]Refers to the dynamic capability of the app to engage users through direct and responsive interactions.

[e]Integrates physical activity into routine daily tasks to reduce perceived barriers and enhance the practicality of exercises.

[f]This component involves the app's capability to synchronize with wearable technology to gather continuous physiological data, which can be used for monitoring health conditions in real time.

[g]Encourages the formation of supportive social networks within the app, enhancing user engagement through community building.

[h]Delivers evidence-based health information and instructional content to improve knowledge and skills related to physical activity.

[i]Uses motivational elements such as web-based badges, achievement unlocking, and progress tracking to enhance motivation and encourage continual app engagement.

[j]Accessibility ensures that products and services are usable by people with various abilities, whereas inclusivity focuses on creating environments that accommodate and welcome diverse individuals across all backgrounds and needs.

[k]An older adult–specific feature is a design element in a product or service, particularly mobile apps, that addresses the physical, cognitive, and social needs of older adults, such as mobility, sensory impairments, chronic conditions, and isolation, to enhance usability, safety, and autonomy.

[l]Presence of component.

[m]PACE: Physical Activity Cardiorespiratory Exercise.

[n]MIA: More In Action.

Features

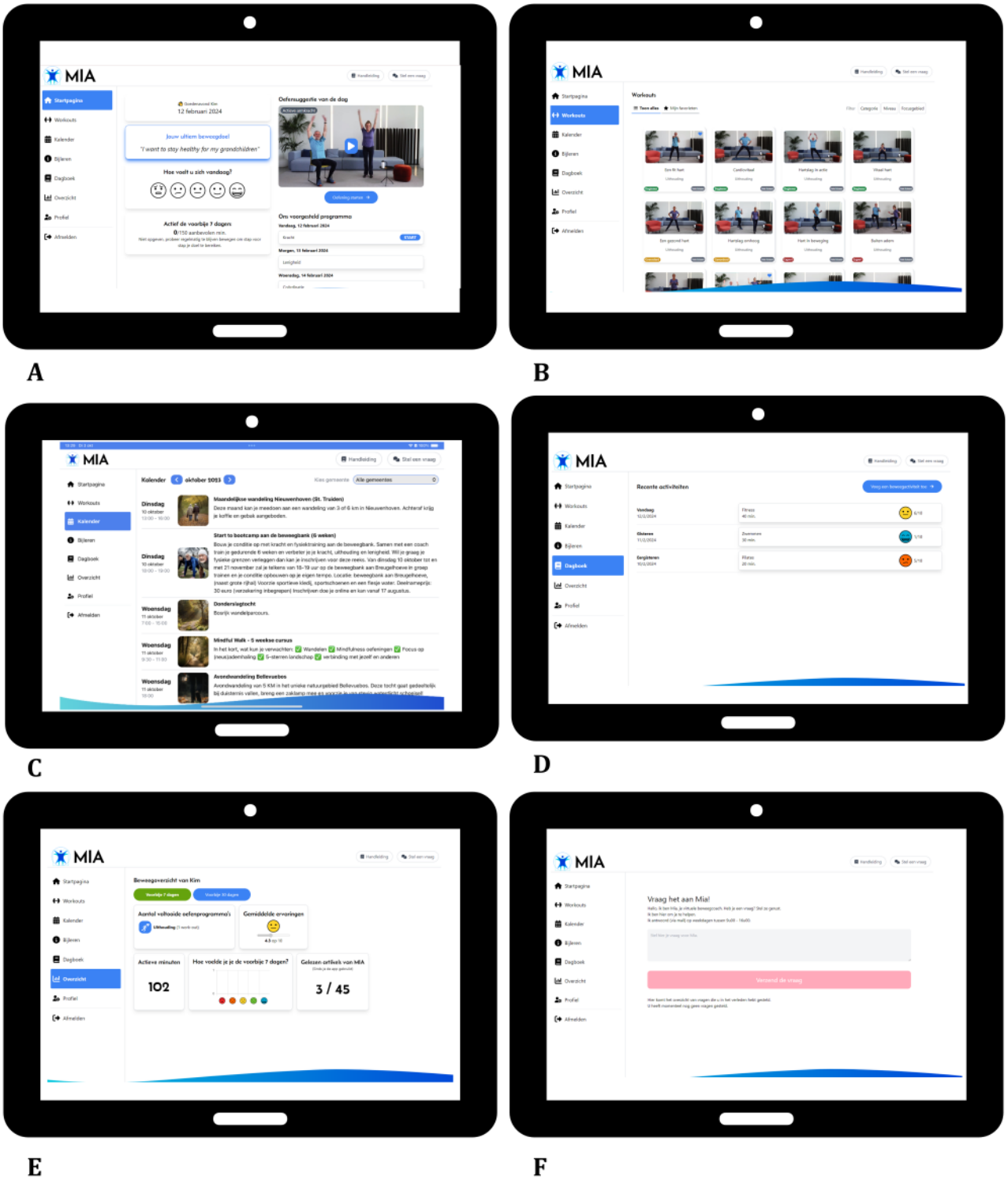
MIA incorporates 7 major features: tailor-made tips, literacy initiatives to enhance awareness about PA and a healthy lifestyle, guided exercise workouts with peers, a community calendar fostering social connections, a personal diary for manual uploads of non–app-based PAs, a progression monitor, and a chatbot named MIA. The chatbot serves as a platform where users can pose questions and receive insights, motivation, and support in return. The current chatbot function is managed manually, with researchers on the development team directly addressing user inquiries. This approach allows for personalized responses and ensures that user concerns are addressed with accuracy and expertise but also for the collection of data that will be used later on to develop an artificial intelligence (AI)–driven chatbot. This AI chatbot would maintain human oversight to ensure the accuracy and relevance of the responses, particularly for complex queries. The app also includes 195 guided workout videos, each featuring 30-second exercise bursts followed by breaks. These exercises cover strength, endurance, coordination, and balance and are demonstrated by peers. Workouts are tailored to 3 skill levels—beginner, advanced, and expert—and range from 10 to 20 minutes. MIA continuously adapts workouts based on user experiences; for example, if an exercise is too challenging, the algorithm adjusts future sessions to ensure that they remain enjoyable and achievable.

User Onboarding and Interaction

Upon initiating MIA, users undergo a preliminary assessment through a questionnaire that informs the customization of their exercise regimen. This personalization is central to MIA's approach, tailoring the app experience to individual needs and preferences.

As shown in Figure 3, the home page presents a personalized PA agenda that adapts daily based on user feedback to align with their physical and mental health status. This adaptive feature is critical for tailoring the experience to the user's day-to-day condition, emphasizing the interconnection between mental well-being and PA.

Figure 3. Screenshots of the More In Action interface: (A) home page displaying a physical activity (PA) agenda, suggested workout videos, and daily tips designed to engage users with personalized fitness guidance; (B) library of workout videos with filters for exercise type, fitness level, and target body areas supporting a tailored workout experience; (C) a community calendar for local PA events; (D) manual logging of nonapp PAs; (E) PA progress monitor; and (F) a web-based coach. MIA: More In Action.



Throughout the intervention period, participants were not required to use the app daily. Instead, following the principles of behavior change techniques, autonomy was emphasized, allowing users to set personalized goals and choose their own frequency of app engagement. This flexibility enabled participants to engage with the app at their preferred rate, from as little as once a week to daily over the course of the 5-week intervention.

Outcome Measures

Quantitative Measures

At Baseline

Demographics and PA levels were first assessed during the baseline measurement. The International Physical Activity Questionnaire–Short Form (IPAQ-SF) [54], the Physical Activity Enjoyment Scale (PACES) [55], and the Digital Health Readiness Questionnaire (DHRQ) [56] were used to measure digital health readiness via a web-based questionnaire using Qualtrics (Qualtrics International Inc). The IPAQ-SF showed moderate validity with accelerometry but had wide limits of agreement, indicating caution for longitudinal use [57,58]. Both the PACES [55,59] and DHRQ [56] demonstrated strong internal consistency and reliability, making them suitable for assessing enjoyment of PA and digital health readiness, respectively.

The 5-Week Postintervention Assessment

To evaluate user satisfaction and experience, 3 established instruments were used: the System Usability Scale (SUS) [60], Net Promoter Score (NPS) [61], and Customer Satisfaction Score (CSAT) [62]. The SUS provides a standardized assessment of usability through 10 items with response options ranging from *Strongly agree* to *Strongly disagree*. Scores of >68 indicate above-average usability. The SUS has demonstrated high internal consistency (Cronbach α=0.74) and test-retest reliability (Pearson correlation coefficient=0.75), supporting its use in evaluating digital health interventions [63]. The NPS measures user loyalty by asking participants to rate their likelihood of recommending the app on a scale from 0 to 10. This metric is calculated by asking older adults the following: *On a scale from 0 to 10, how likely are you to recommend MIA-app to a friend or peer?* Respondents are classified into 3 categories: promoters (those who provide a score of 9-10), passives (those who provide a score of 7-8), and detractors (those who provide a score of 0-6). The NPS, derived by subtracting the percentage of detractors from that of promoters, offers a straightforward approach to understanding user loyalty and referral potential [62].

The CSAT quantifies user satisfaction using a 10-point scale, capturing immediate feedback on the app experience. App-specific concerns (user-friendliness, layout, and utility) were also evaluated across multiple categories, with ratings ranging from 1 to 5 (1-2 indicating dissatisfaction, 3 indicating neutrality, and 4-5 indicating satisfaction).

Microsoft Power BI (Microsoft Corp) was used to analyze participant behavior during app use to provide insights into app use analytics and user engagement patterns [64]. Data on metrics such as workout frequency, session duration, and instances of premature workout termination were collected using integrated tracking tools such as Plausible Analytics. These data were imported into Microsoft Power BI for further analysis. Web-based dashboards and visualizations, including bar charts, were created to represent key engagement metrics and user navigation patterns.

Qualitative Measures

At Baseline

A *concurrent think-aloud* approach allowed participants to voice their thoughts and actions while using the app throughout installation and use [65,66]. Usability and enhancement suggestions were supplied to the researchers in real time [67-69]. Multimedia Appendix 3 details the think-aloud protocol, with an average procedure time of 20 (SD 6.2) minutes. These sessions were conducted in our laboratory at PXL University of Applied Sciences and Arts.

The 5-Week Postintervention Assessment

Qualitative data from focus groups were also collected to better understand app users' experiences. Following the 5-week trial of the app, the participants were invited to take part in focus groups in our laboratory at PXL University of Applied Sciences and Arts, each consisting of 6 older adults. Each focus group session lasted 2 hours and aimed to explore participants' perceptions of the app along with their reflections and suggestions for improvement. The sessions covered all app features. The focus groups were facilitated by experienced researchers (JR and KD) who are specialized in qualitative research. Multimedia Appendix 4 contains the interview guide.

Data Analysis

Quantitative Analysis

The quantitative analysis used mostly descriptive statistics to examine the data. Frequency distributions were also used to observe the occurrence of data values or categories, offering a detailed view of data distribution across outcomes. In addition, Microsoft Power BI analytics played a crucial role in visualizing and interpreting the data. Microsoft Power BI allowed for dynamic dashboards and visualizations, making it easier to explore trends, patterns, and relationships within the data. By integrating various data sources, it enabled the creation of charts, graphs, and tables that provided real-time insights into participant behavior and app use.

Qualitative Analysis

The think-aloud method and focus groups were audio recorded with participants' consent for the qualitative component. The transcriptions of the focus groups were carried out verbatim, followed by a rigorous content analysis using the NVivo software (QSR International) [70]. Themes were systematically coded based on the guidelines by Braun and Clarke [71]. The coding process was independently carried out by 2 experienced researchers (KD and JR), ensuring reliability and reducing the potential for bias in theme interpretation.

Integration of Qualitative and Quantitative Results

The Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework was used to integrate both the qualitative and quantitative results of the study, providing a structured analysis of the potential for sustainable implementation of the MIA app [72]. Recognized for its utility in evaluating complex technological health interventions, the NASSS framework allows for a systematic examination of key factors influencing the deployment and long-term sustainability

of digital health technologies. The analysis encompassed the framework's 7 domains: the condition being addressed (physical inactivity among older adults), the technological characteristics of the MIA app (including usability and technical robustness), the value proposition (perceived benefits to end users and stakeholders), the adopter system (end users and supporters), the organizational context (integration within health care systems), the wider sociopolitical environment (regulatory and cultural considerations), and the processes of embedding and adapting the technology over time.

Results

Demographic Characteristics

A total of 30 participants (mean age 70.3, SD 4.8 y; n=17, 57% female) were included in this study. Participants had different educational backgrounds; the largest proportion (15/30, 50%) had completed higher education. Their level of digital literacy was commendable, displaying a mean score of 59.6 (SD 8.8) out of 75 on the DHRQ. Notably, 80% (24/30) of the participants acknowledged the affirmative impact of digitalization on health, whereas 73% (22/30) actively engaged with social media, among whom 36% (8/22) used it daily, 23% (5/22) used it often, 23% (5/22) used it occasionally, and 14% (3/22) used it rarely. Furthermore, 67% (20/30) of the participants sought health-related information on the web, whereas all 30 participants reported internet use, with 27 (90%) of them accessing it daily, and 17 (57%) owned a smartwatch. When it came to monitoring PA through their smartphones' health apps, 73% (22/30) of the participants checked their step count daily, 13% (4/30) did so frequently, 3% (1/30) did so occasionally, 3% (1/30) did so rarely, and 10% (3/30) never did. Regarding health-related app use, 17% (5/30) of the participants actively used this type of apps.

The IPAQ-SF revealed variability in PA levels among individuals. The median activity levels indicated moderate engagement in walking (1386, IQR 284.8-2178 metabolic equivalent of task min/wk) and lighter involvement in vigorous activities (median 0, IQR 0-1830), suggesting a skew toward lower-intensity exercises. Moreover, it was found that 10% (3/30) of the participants were classified as inactive, 40% (12/30) fell into the category of minimal activity, and a significant 50% (15/30) were categorized as vigorously active. High SDs across all categories underscored a wide range of PA levels, from minimal to extremely active, reflecting the diverse nature of physical engagement in this sample.

The PACES results revealed that participants, on average, reported a moderate to high level of enjoyment during PA, with a mean score of 107 [55]. Considering that the PACES scale ranges from 18 to 126, where higher values signify greater enjoyment, the mean score of 107 suggests that participants' enjoyment levels were significantly skewed toward the upper end of the scale. This positioning implies a generally positive perception of PA among the participants. The complete sociodemographic characteristics of the participants are presented in Table 3.

Table 3. Participant characteristics (N=30).

Variable	Values
Age (y), mean (SD)	70.3 (4.8)
Sex (female), n (%)	17 (57)
Marital status, n (%)	
Single	2 (7)
Living together	1 (3)
Married	21 (70)
Divorced	4 (13)
Widowed	2 (7)
Educational level, n (%)	
Primary school	1 (3)
Middle school	6 (20)
University of applied sciences	15 (50)
University	8 (27)
Fall incidence (yes), n (%)	3 (10)
Digital literacy score (DHRQ [a] ; out of 75), mean (SD)	59.6 (8.8)
Use (out of 20)	16.3 (3.1)
Skills (out of 25)	21.3 (3.2)
Literacy (out of 15)	12.3 (2.2)
Health literacy (out of 15)	9.8 (3.2)
Learnability (out of 25)	20.7 (2.5)
Physical activity level (IPAQ - SF [b]), median (IQR)	
Total	3273 (1345.5-3873)
Walking	1386 (284.8-2178)
Moderate	1020 (310-2070)
Vigorous	0 (0-1830)

[a]DHRQ: Digital Health Readiness Questionnaire.

[b]IPAQ-SF: International Physical Activity Questionnaire–Short Form.

Quantitative Analysis

Acceptability

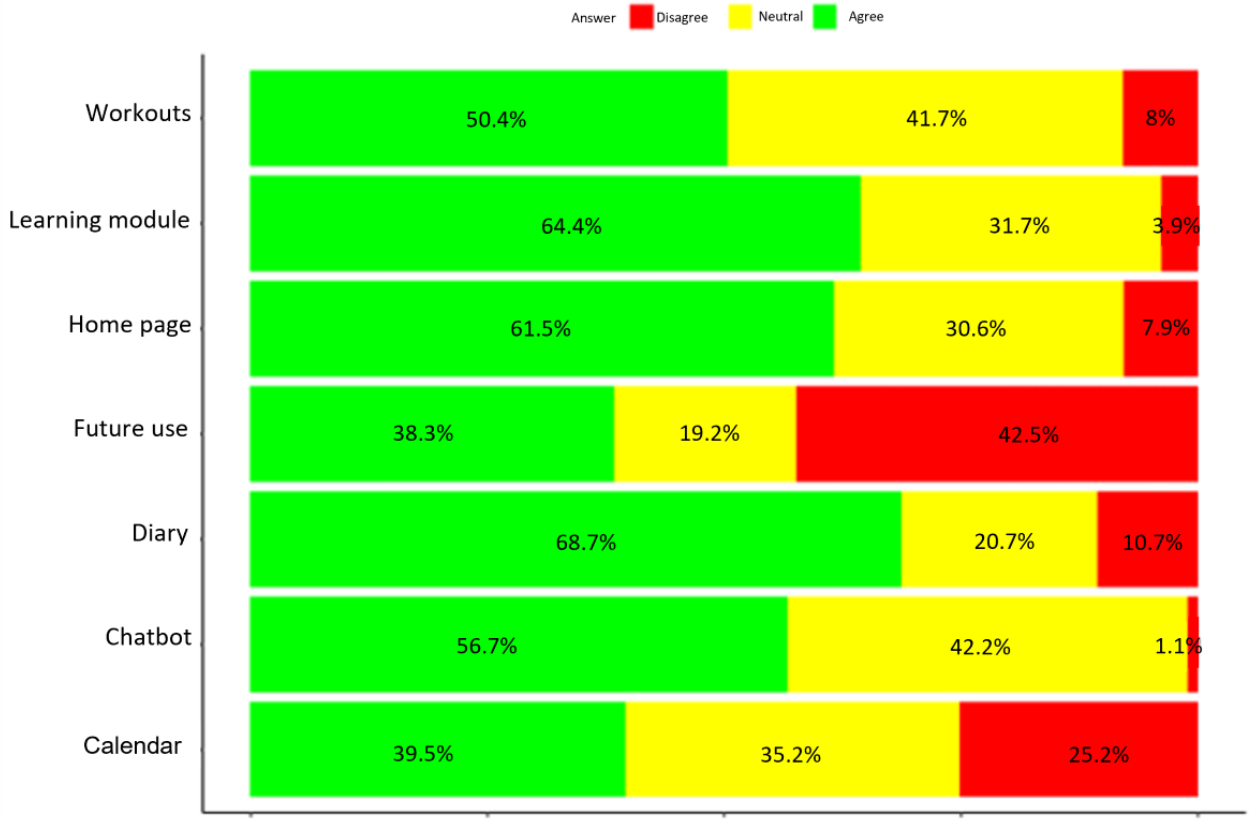
First, the general acceptability was accessed using the 3 main indicators, as shown in Figure 4.

Figure 4. Results of the System Usability Scale, Customer Satisfaction Score, and Net Promoter Score (NPS). MIA: More In Action.

According to the SUS, the mHealth app was considered *acceptably good*, with a mean score of 77.4 (SD 14.3). The results of the CSAT revealed that 47% (14/30) of the participants indicated that they were *very satisfied*, whereas 40% (12/30) reported being *satisfied*. Another 10% (3/30) had a neutral response, and only 3% (1/30) were *dissatisfied*. None of the participants provided a *very dissatisfied* score. This distribution led to a CSAT score of 86.6%, indicating a high satisfaction level among older adults using the MIA app. The NPS was used to assess customer satisfaction with and loyalty and enthusiasm toward the app. A total of 50% (15/30) of the participants were *promoters*, indicating that they would recommend the MIA app to peers and family. In addition, 33% (10/30) were *passives*, reflecting moderate satisfaction without strong advocacy, whereas 17% (5/30) were *detractors*, indicating that they were unlikely to recommend the app. The resulting NPS is calculated by subtracting the percentage of detractors from the percentage of promoters, yielding a total NPS of 33.34.

Furthermore, the specific acceptability of the different features was assessed. The quantitative results also guided the qualitative part of the study, where focus groups were conducted to dive deeper into participants' reflections and gather additional insights. The aggregated results are presented in Figure 5. The complete results per category are presented in Multimedia Appendix 5.

Figure 5. User satisfaction and engagement—categorized ratings of app features and interface (1-2: dissatisfaction [red]; 3: neutrality [yellow]; 4-5: satisfaction [green]).



User Behavioral Patterns

To track user navigation and behavior within the app, Microsoft Power BI analytics were collected on the following features: workout videos, learning modules, community calendar, manual diary, progression monitor, and chatbot. Visual representations of the Microsoft Power BI analytics are available in Multimedia Appendix 6.

Workout Videos

A total of 585 workout videos were analyzed, highlighting general satisfaction with a median satisfaction score of 4 (IQR 3-4; scale of 1-5) and physical exertion rated at a median of 5 (IQR 5-7; scale of 0-10). Users engaged for a median duration of 14 (IQR 0-34) minutes per session, with use times ranging from 0 to 34 minutes. Strength workouts were most common, comprising 40% (234/585) of the sessions, followed by endurance, flexibility, and balance workouts. Most workouts (429/585, 73.3%) were aimed at beginners, with only 5% (29/585) targeting expert levels, reflecting the app's focus on older adults. Notably, 77% (23/30) of users expressed no preference for a specific type of exercise, suggesting a wide acceptance of the available workout options.

Learning Modules

The distribution of the learning modules comprised 255 read articles across 6 different topics: mental health, social well-being, physical well-being, nutrition, sleep, and risk of falling. Physical well-being was the most frequently engaged with topic, with 26.3% (67/255) of the total read articles. Nutrition followed with 17.3% (44/255) of the read articles, whereas mental health comprised 16.9% (43/255) of the read articles.

Community Calendar

Only 3 activities were added to the participants' community agenda, all of which were organized walks.

Manual Entries in Diary

The analysis of 545 external activities showed a median *general feeling after exercise* score of 4 (IQR 4-4) on a scale from 2 to 5, indicating generally positive responses. The range of perceived exertion scores was broader, with a median of 5 (IQR 5-7) indicating a balance between lower and higher levels of perceived effort. Walking emerged as the most popular activity (210/545, 38.5% of external activities), followed by cycling with electric assistance (78/545, 14.3%); cycling without electric assistance (66/545, 12.1%); and other activities such as padel, tennis, strength training, and running. The duration of these activities varied significantly, with a median time of 60 (IQR 40-90) minutes ranging from 5 to 400 minutes, indicating diverse engagement levels. The data revealed that a quarter of the participants (8/30, 27%) engaged in activities for <40 minutes, whereas a significant portion spent ≥90 minutes.

Chatbot

During the 5-week period, the chatbot was consulted 66 times. Most of the interactions involved feedback, accounting for 48% (32/66) of instances. Communication regarding exercise differentiation occurred 15% (10/66) of times, whereas

medical-related queries were raised 5% (3/66) of times. Technical issues led to 9% (6/66) of consultations, and motivational support prompted 17% (11/66) of interactions. The remaining 6% (4/66) of consultations fell into other categories.

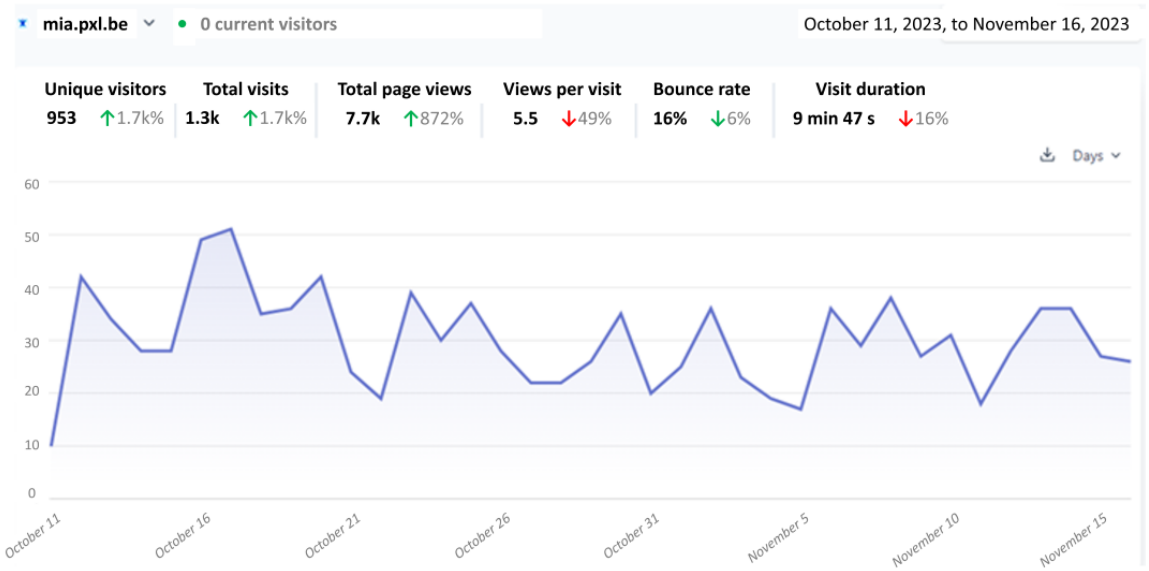
Adherence

During the 5-week intervention, app adherence was measured by tracking both the frequency of use and interaction patterns. Rather than requiring daily use, the study sought to assess whether participants engaged with the app according to the personalized goals and self-determined use frequencies they had set. The intervention, guided by self-determination theory [33], prioritized autonomy, enabling participants to select their own goals and decide how often to use the app. Engagement levels varied, with participants using the app between 3 and 7 times per week, with a median of 5 (IQR 4-6) times per week. Adherence rates averaged 135%, with individual rates ranging from 21% to 271%, reflecting the diverse levels of user interaction throughout the intervention.

On average, participants engaged with the app 32.88 (SD 14.06) times, ranging from 11 to 59 times used, which translates to an approximate daily use rate of 0.94 (SD 0.40) times. This indicates near-daily use of the app among the study cohort.

Despite the considerable variability in workout interactions within the app, with a median workout frequency of 17 sessions (IQR 11.2-28.5), motivation remained notably stable throughout the study, with some dips during the weekends. Participant engagement and interaction patterns throughout the 5-week intervention are visualized in Figure 6. Daily engagement varied between 0.66 and 3 sessions per day; however, there were no discernible fluctuations in adherence over time. Unlike the common pattern of an initial surge in activity followed by a dip and potential recovery, participants demonstrated consistent interaction with the app across the 5-week intervention.

Figure 6. Participant engagement and interaction patterns over the 5-week intervention. The y-axis shows the number of visits to the More In Action app.



Qualitative Analysis

Think-Aloud Method

The thematic analysis of the *think aloud* data resulted in the identification of 3 major key themes presented in Textbox 2. The other results and functionality are briefly described in the following sections.

Participants began interacting with the MIA app by opening it and completing their profiles. The welcome screen was inviting and easy to navigate, but some participants encountered difficulties with the profile questionnaire, highlighting a need for clearer guidance. When exploring the home screen, the interface was straightforward, with large icons and a simple layout. However, some features such as mood indicators were confusing, indicating a need for better explanations. The workout section was well received for its variety and suitability for different fitness levels, enhancing user experience. However, navigating the calendar feature proved challenging for some, with issues in adding or removing items pointing to a need for a more intuitive design. The learning module page was appreciated for its informative content, although some noted a lack of retirement-related topics, suggesting room for expansion. The diary functionality faced some hurdles as users found it difficult to add past activities, needing more flexible controls. Visually, the app was praised for its appealing design with bright colors and clear typography, making it esthetically pleasing and not overly complex. However, specific challenges such as difficulties with the profile questionnaire and mood indicators revealed that clarity in guidance is also crucial. Thus, while the app's visual appeal contributes to its overall usability, the complexity of certain features points to the need for balancing esthetics with intuitive guidance.

Textbox 2. Key themes of the think-aloud protocol.

Usability and accessibility

- The participants' experiences reveal that usability and accessibility are crucial for a successful app. Challenges with completing profiles, navigating calendars, and adding diary entries highlight areas in which the app could improve. Ensuring that the app is intuitive and easy to use can lead to a better user experience.

User interface and design

- The overall design and layout of the app play a significant role in user satisfaction. The positive feedback on the app's simplicity, large icons, and bright colors shows that a visually appealing and straightforward design enhances usability. This theme indicates that maintaining a clean and user-friendly interface is essential.

Feature functionality and clarity

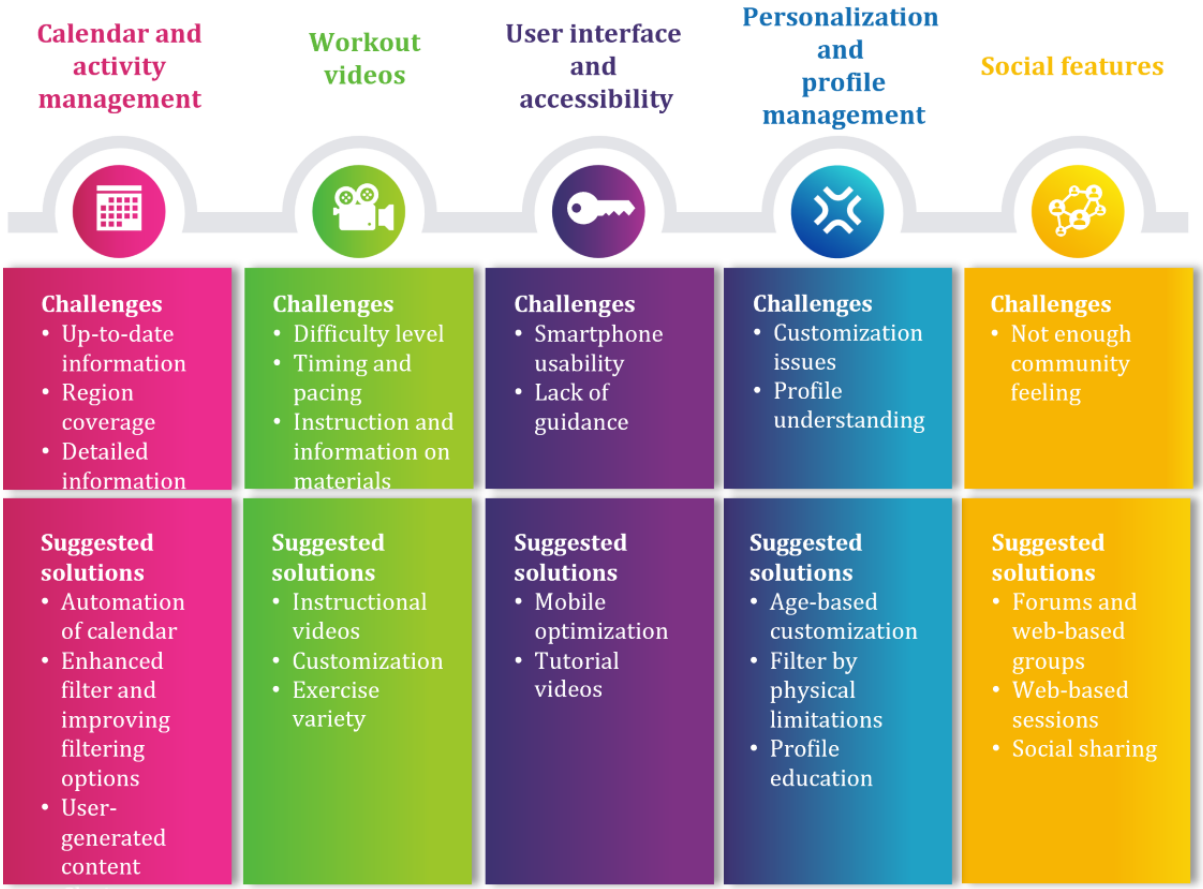
- The confusion regarding certain features, such as the mood indicators and the calendar, suggests that functionality and clarity are important. Ensuring that each feature is well explained and operates smoothly can lead to a more satisfying user experience. This theme points to the need for clear communication and improved design in specific areas of the app.

Focus Groups

Overview

The transcripts of these focus groups led to identifying 5 key themes, presented in Figure 7: calendar and activity management, workout videos, user interface and accessibility, personalization and profile management, and social features. Participants valued the community calendar for fostering social connections, but engagement in activities was low. Tailored workout videos were appreciated for their adaptability, although more flexibility in goal setting and detailed instructions were desired. Feedback highlighted the intuitive interface but suggested improvements in flexibility, readability, and tracking. Personalization features such as the chatbot were appreciated, although users wanted better clarity in activity plan customization. Social features, including a buddy system, were seen as ways to boost user interaction and engagement. Each of these themes is further detailed in the following subsections.

Figure 7. Key themes of the postassessment focus groups.



Calendar and Activity Management

Participants valued the community calendar for enhancing social connections but reported limited participation in the listed activities. A participant aged 68 years highlighted the following:

> *The agenda is there because of the social connection with others. I really liked it, although I didn't go to an organized activity. You should definitely keep it in there.*

Others suggested that participation might increase with better visibility and advertising of the activities, along with reminders and notifications to prompt more active involvement.

Workout Videos

The tailored workout plans and videos were popular among users. A participant aged 67 years wanted greater control over his workout goals, saying the following:

> *The target of 150 minutes per week is good, but I would like to choose my target minutes myself, for instance, 500 minutes.*

This comment indicated that some users prefer more flexibility in setting personal goals. Another participant aged 72 years commented the following:

> *The algorithm when a workout is too challenging works well; the next video suggestion is indeed tailored to my level. I tested it to see if it would work.*

This adaptability was valued, but users pointed out that the video instructions could sometimes be too quick to follow. Another participant aged 70 years who was already active and doing a lot of cardiovascular exercise mentioned the following—"The app and the exercises made me realize I need to work on my strength as well"—suggesting that the app can broaden users' fitness horizons.

Participants also suggested improvements such as being informed about required materials in advance. A participant aged 69 years noted the following:

> *It would be good to know which materials to use in advance.*

This feedback suggests that clearer preparation instructions could enhance user experience.

User Interface and Accessibility

Various aspects of the app's interface and accessibility were highlighted, noting strengths and areas for improvement. The calendar feature was appreciated for fostering social interactions, although it lacked sufficient engagement tools. Users found the manual diary and progression monitor helpful for tracking activities. Still, they criticized their limited functionality, particularly the inability to edit past entries and the short view range of 7 days. Suggestions included enhancing flexibility and extending tracking capabilities to improve usability.

The *Learning* section received mixed reviews; while the content was engaging, users suggested features such as checkboxes to track articles read, indicating a need for more interaction and relevance to older adults. Although the app's layout was praised for its esthetic and ease of navigation, calls for better readability and customization were prominent.

Feedback on the chatbot was mixed, with some users finding it motivating whereas others saw room for personalization. General usability issues related to smartphone optimization and intuitive navigation were raised but reportedly resolved quickly, demonstrating effective technical support. While the app was generally well received, users desired more robust features and personalization to enhance their experience.

Personalization and Profile Management

Users appreciated features that allowed for personalization, such as the chatbot and manual diary, recognizing their value despite some limitations. A user aged 68 years praised the diary feature, stating that "The diary was great. It is really good to track your activities," but also recommended enhancements, suggesting that "It would be nice if the diary could go back in the past, more than now." In addition, it became apparent that users were unaware that the initial registration questions were intended to personalize their plans as a user stated the following:

> *I didn't know we could adjust the information in the PA plan.*

This misunderstanding also led to them not expecting the system to adapt over time.

Nonetheless, several participants acknowledged the app's added value in promoting PA. A participant aged 66 years noted the following:

> *Getting the push notifications motivated me; it was like a digital motivator in being more active.*

Social Features

The social features such as the community calendar were highlighted as significant opportunities to enhance user engagement. A participant aged 66 years suggested a buddy system to increase interaction:

> *Perhaps a buddy system where I can see my friend and we can motivate each other.*

Suggestions for Improvement

Multimedia Appendix 7 summarizes the suggestions and iterations from the focus groups, offering insights into lessons learned during the development process.

Integration of Quantitative and Qualitative Analysis

The NASSS framework [72] provided a structured approach to integrating the quantitative and qualitative results, revealing barriers to and facilitators of adoption and the interactions influencing scalability and sustainability. It enabled comparisons among user experiences, usability scores (eg, SUS and CSAT), and app interaction data while incorporating qualitative insights from focus group feedback on usability challenges and user perceptions. This enabled us to identify alignment and discrepancies between users' perceived value of the MIA app and their actual app interactions, as well as broader contextual factors influencing sustained use. Figure 8 visualizes this analysis, emphasizing how the NASSS framework bridges the gap between numerical data and user narratives.

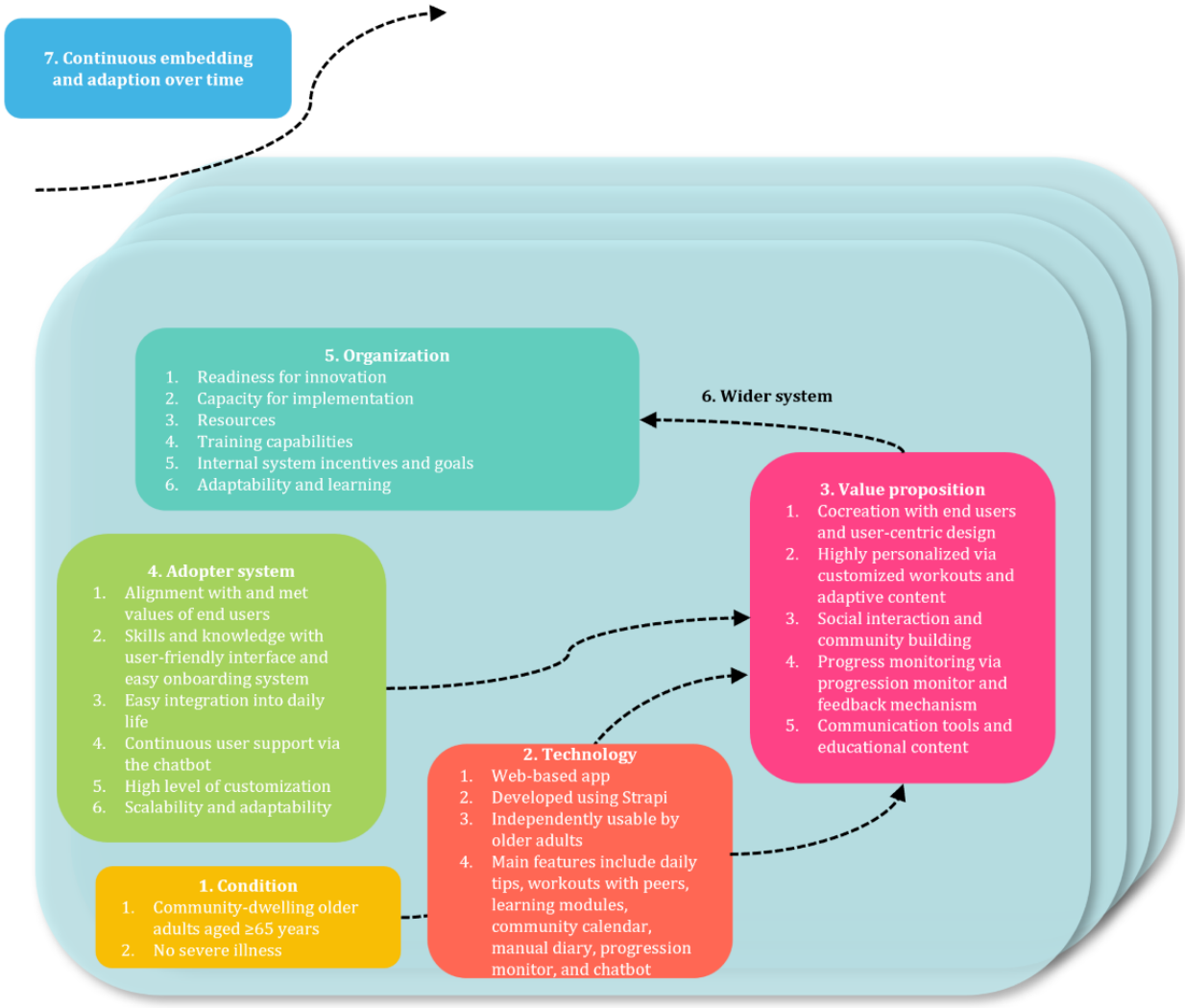
Addressing physical inactivity among older adults is challenging due to varying digital literacy, physical abilities, and health conditions. A key facilitator was the app's customization, allowing users to adjust exercise recommendations, leading to positive SUS and CSAT scores. The app's technical stability and responsive support further encouraged adoption among digitally literate users.

However, barriers such as navigation difficulties and insufficient guidance on features such as profile setup and calendar use limited engagement, with 30% (9/30) of the users classified as passive and 17% (5/30) classified as detractors in the NPS, indicating a risk of user churn. While users appreciated tailored workouts and real-time feedback, some sought greater control over goal setting, suggesting the need for enhanced personalization.

The NASSS analysis identified opportunities to boost engagement, including social features such as a buddy system to align with older adults' social preferences. Integration with health care systems could facilitate broader adoption if endorsed by providers. Alignment with health policies and partnerships may also expand the app's reach.

Despite these strengths, challenges such as technological intimidation, market competition, age-related declines, and data privacy concerns pose risks to long-term adoption. Adherence to ethical standards mitigated some concerns. Sustained engagement depended on the app's adaptability to user feedback, requiring continuous updates and technical support to maintain user interest. Detailed information is available in Multimedia Appendix 8.

Figure 8. Nonadoption, Abandonment, Scale-up, Spread, and Sustainability framework adapted for the feasibility study of the More In Action app.



Discussion

Principal Findings

In recent years, the intersection of aging and technology has gained recognition as a vital factor in enhancing quality of life for older adults [73]. Technology holds substantial potential to improve everyday living for this demographic, yet adoption barriers remain a significant concern as outcomes often rely on user-driven interactions, which can vary significantly [10,34,74,75]. Effective technology design for older users is best achieved through *user involvement and cocreation*, a

strategy that leverages the unique insights of users, who are experts in their own experiences [19,20,61,76,77]. This participatory approach not only results in products that better meet older adults' specific needs, desires, and challenges but also counters age-related stereotypes in gerontechnology [78]. MIA exemplifies this by having been developed through a collaborative cocreation process with older adults. However, to achieve acceptable technologies for older adults, an in-depth understanding of their acceptability, feasibility, and usability is necessary [79]. Against this backdrop, the main aim of this study was to thoroughly explore the acceptability, feasibility, and usability of the MIA app in promoting PA and encouraging sustained, active, and healthy behaviors among its users.

The app received positive feedback, achieving an SUS score of 77.4 and a high CSAT rate of 86.6%. The NPS stood at 33.34, indicating a good level of customer loyalty, with half (15/30, 50%) of the participants categorized as promoters. Feedback on the app was positive for its clarity and engagement. In addition, feasibility and adherence levels were comparable to those observed in other apps evaluated. For example, StandingTall showed acceptable usability, with exercise adherence improving over time [35], whereas Physitrack reported a high study retention rate of 95% and an adherence to prescribed exercises of 84% [40]. Similarly, Fit for All [43] recorded an adherence rate of 82%, supporting the app's potential in facilitating engagement and promoting PA among users. The workout videos on the MIA app received a commendation, although some participants suggested improvements. The learning module and diary functions were praised for their utility, yet some users noted challenges with navigation and desired the ability to edit past diary entries. The progression tracker and chatbot were also seen as valuable, although there was a call for more detailed tracking and enhanced personalization. Analysis showed strong engagement with the workout videos, especially strength-focused sessions, and high satisfaction levels with these workouts. Insights from the think-aloud protocols and focus groups highlighted the need for clearer guidance on profile management, workout customization, and calendar functionality. Participants also expressed a desire for more engaging social features, such as a buddy system. In general, the app was well received for its educational content and tracking features. However, there were recommendations for improving the user interface, enhancing social connectivity, and further personalizing the content.

There are 3 reasons why the MIA app offers more value compared to other alternatives (Table 2). First, the development and design process of the MIA app was based on the design thinking method and user-centered approach, guided by older adults at every stage. Their input was pivotal, ensuring that the app's features were tailored to their specific needs and preferences. Moreover, most existing apps in this domain predominantly cater to the younger segment of the older population, typically those aged 50 to 55 years [10,12,13,80-82]. This demographic focus leaves individuals aged ≥65 years underrepresented. The MIA app addresses this gap by concentrating research and design efforts on the age group of ≥65 years, thereby providing a more inclusive and representative technological solution for older adults.

Second, the MIA app incorporates behavior change theories such as the BCW [29], self-determination theory [33], and self-identification theory [30] to enhance user engagement.

Finally, the MIA app is engineered to dynamically adapt to the diverse capabilities and fluctuating motivations of its users [83,84]. It features an extensive repertoire of 195 distinct exercises encompassing various PA domains such as strength, coordination, balance, flexibility, and endurance, as well as incorporating activities pertinent to daily living. During the feasibility study, sending push notifications with personalized health messages was linked to increased user engagement. In addition, research indicates that users are more likely to interact with the app within 24 hours when push notifications are sent at midday on weekends [85-87]. Personalization is achieved by accounting for the inherently dynamic nature of motivation, which is known to vary not only from day to day but also within a single day [88]. To systematically capture these fluctuations, the app uses the technique of ecological momentary assessment [89]. This approach involves the intensive and recurrent collection of data on an individual's behavior and motivation in a real-time environment [90-92]. Moreover, the app continuously refines its exercise recommendations based on user interactions and feedback. If an exercise is deemed excessively challenging, the algorithm records this, adjusting future workouts to better match the user's current abilities and psychological state. This ensures that subsequent sessions are both enjoyable and attainable, fostering ongoing engagement and facilitating progressive improvement.

To ensure long-term retention and motivation, new videos and activities are regularly added to the app. Future updates to the MIA app will emphasize social features and gamification, including a buddy system for users to connect with exercise partners for mutual support, as well as earning badges, completing challenges, and participating in leaderboards for friendly competition. These elements, along with personalized workouts, progress tracking, content refreshers, and timely reminders, aim to enhance user engagement and sustained adherence.

Strengths and Limitations

The results of this study should be interpreted in light of some limitations. First, the participant profile may not fully represent the broader older adult population. The individuals involved in this feasibility study were notably well educated and digitally literate, with nearly 60% (17/30, 57%) owning smartwatches. This contrasts with broader trends identified in a recent study [93], which found that only 25% of those aged ≥65 years use smartwatches or health apps, a number that declines to 16% among those aged >75 years. In our study, most participants already owned smartphones and were more acquainted with mobile technology. Future studies could explore the perceptions and willingness to adopt this mHealth app among older adults less familiar with such technology.

In addition, the participants who volunteered for this study might have been inherently more motivated and interested in PA than the general older adult population. This could limit the generalizability of the findings as their feedback might not reflect the perspectives of a broader, less technologically adept

audience. Involving representatives of the older people demographic was very challenging because older adults are an extremely heterogeneous group with highly varied characteristics and needs who use, modify, and interact with technologies in rather diverse ways [94]. To mitigate the limitations of this study and enhance its generalizability, several strategies need to be implemented in the future. This involves targeting individuals who are less educated and less familiar with digital technology and those who do not own smart devices. Collaborating with general practitioners can help reach this more varied participant pool.

In addition, focus groups with older adults can sometimes result in conflicting opinions regarding an app's design and features, which proves that a one-size-fits-all approach does not work for this population. Fortunately, the opinions of the older adults were well studied and were considered as much as possible during the iteration process after the feasibility study. To address the issue of conflicting opinions in focus groups, it would be beneficial to implement a more personalized approach in future studies [95,96]. This could include individual interviews that can capture a wider range of perspectives and experiences. This would allow researchers to explore deeper insights and accommodate each participant's unique needs and preferences, moving beyond a one-size-fits-all solution.

To enhance both the reliability and validity of future studies, it is recommended to incorporate randomized controlled designs that include blinding procedures. This approach will help minimize bias by ensuring that assessors do not know which participants have received the intervention, thereby providing a more objective assessment of the outcomes [97,98]. In addition, conducting longitudinal assessments could yield a richer understanding of the intervention's long-term attrition rates, further strengthening the study's findings and their applicability across different settings and populations [99,100].

Finally, unrealistic expectations surrounding an mHealth app could pose a challenge. Users might expect immediate and significant improvements, failing to recognize that behavior change is typically a gradual process [24]. Therefore, it is crucial to underscore that the app is designed to be a supportive aid in the journey toward change, not a quick fix. To address this issue, managing expectations proactively is essential. Effective communication and educational initiatives are key to establishing a realistic understanding of what the app can and cannot do. By clearly outlining the app's functionalities and limitations, users can be better prepared for their experience. This approach helps maintain sustained engagement and maximizes the app's potential as a tool for positive transformation.

Despite these limitations, this study's primary strength lay in its adoption of a mixed methods approach, which facilitated a comprehensive understanding of the app's feasibility. The methodology combined surveys using validated tools, interviews, focus groups, and Microsoft Power BI analytics. This robust approach enabled us to construct a detailed profile of user behavior within the app. Through the systematic analysis of interaction data, we identified the features that garnered the most engagement and pinpointed areas for potential enhancement to improve user experience and retention. The insights derived from this extensive data analysis are fundamental in shaping the future development strategies of the app, ensuring that it aligns more closely with user needs and preferences. Furthermore, by using these diverse research techniques, we were able to gather rich data on participants' experiences concerning the feasibility of the MIA app. This collaborative methodology proved especially beneficial, fostering meaningful dialogues between participants and researchers. These discussions yielded valuable perspectives on potential improvements to the mHealth app.

Future Use and Implications in the Field

The MIA app is currently free and open source. However, to ensure its long-term viability, partnerships with health care insurers and other market players are being explored. A willingness-to-pay analysis [101], detailed in Multimedia Appendix 9, indicated that end users are willing to pay €4.50 to €7 (US $4.66 to US $7.25 at a conversion rate of €1=US $1.04) per month for access to the app. This willingness to pay suggests a perceived value of the app beyond its initial free access, which is further supported by user behavior. Specifically, after 6 months, 43% (13/30) of the participants continued using the app daily without being prompted.

Implementing MIA in real-world settings involves integrating it into health care systems, community programs, and support networks, positioning it as a preventive tool. Partnerships with health care providers and insurers can promote MIA's role while community centers leverage its social features to encourage group activities.

As digital health technologies continue to evolve, the MIA app needs to take several opportunities into account to expand its applicability and feasibility. This offers a range of possibilities for the field.

Tailored user experience remains a very important subject to keep motivating older adults by providing relevant and achievable goals that cater to their individual fitness levels and health conditions [87,102,103]. Enhanced just-in-time interventions enabled by the app's real-time data capabilities through ecological momentary assessment could revolutionize preventive health measures [104,105]. This merges with the potential for integration with wearable technologies, promising a more holistic approach to health monitoring that could improve predictive health interventions for older adults [106-108]. Moreover, integration with wearable devices can enhance the MIA app's functionality by providing real-time, accurate data on PA levels and health metrics. This integration can promise a more holistic view of a user's health and a more precise adjustment of their activity recommendations [50].

As social engagement appeared to be a critical component in maintaining motivation for PA among older adults, the MIA app could include more robust social networking features such as support groups, cooperative challenges, and shared fitness goals [30,109,110].

Another important evolution to consider is the evolution of advanced predictive analytics [111-113]. With advancements in AI and machine learning, the app could incorporate predictive

analytics to forecast potential health risks based on user activity and health data. This feature could alert users and health care providers to potential health issues before they become severe, facilitating timely intervention.

In addition, the MIA app's potential expansion to provide targeted support for informal caregivers, especially those managing care for partners with dementia, offers a pathway to significantly alleviate caregiver burden [114,115]. This feature could become increasingly vital during crisis situations such as pandemics, where the app's adaptability could ensure continuous support for physical and mental health under restrictive conditions [116-119]. Furthermore, by enabling older adults to exercise independently, the app could empower health care professionals by reducing the frequency of in-person checkups, thereby optimizing health care resources [120].

Finally, a broader health integration more closely with health care systems allowing for a smoother exchange of information between the app and health care providers could be a great opportunity [121]. This integration could enable the development of personalized health care plans based on the app's insights, enhancing the overall health care experience by keeping physicians informed and engaged in their patients' lifestyle changes.

Conclusions

This study highlights the potential in merging aging with technology to enhance quality of life and prevention through promotion of PA. Despite some limitations, the app received positive feedback for its usability and customer satisfaction. This underscores the value of involving users in the design process, adhering to a cocreation model that caters to their specific needs and counters age-related stereotypes in technology design. However, the insights gathered suggest a need for broader inclusivity in future studies, targeting less technologically savvy older adults to improve generalizability.

The analysis revealed strong engagement with specific app features (eg, workout videos) and highlighted areas for improvement, such as user interface and social connectivity enhancements. Longitudinal studies and ongoing iterations informed by user feedback will be crucial in refining these features. Managing expectations is also essential as technology adoption among older adults often requires recognizing that behavior change is gradual. Future directions include integrating the app into health care systems to tailor health plans more precisely to individual needs and expanding the app's functionality using predictive analytics to pre-empt health issues. Ultimately, by continuing to evolve and adapt to user feedback and technological advancements, technology such as the MIA app can significantly contribute to promoting sustained, active, and healthy behaviors among older adults, demonstrating the profound impact of well-designed gerontechnology.

Acknowledgments

This work was supported by PXL University of Applied Sciences and Arts (2/DWO/2021/HC/P133) and Flanders Innovation and Entrepreneurship (2/DWO/2022/HC/VL041).

Authors' Contributions

KD developed the initial concept for the project and was responsible for acquiring funding. KD undertook the data analysis and curated the necessary resources while also creating visual representations of the data. KD drafted the original manuscript and led the editing and revision process. JR and KQ both contributed to the investigation process and assisted in editing and reviewing the manuscript. SV was similarly involved in the investigation and provided critical review of the manuscript drafts. JB, AS, and DH provided feedback on the manuscript. BB, aside from assisting in manuscript revision, developed the methodological framework of the study and supervised the research process, ensuring adherence to academic standards and contributing significantly to the original draft. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The full informed consent form used to secure participant consent for the study. The form is organized into the following sections: study overview, confidentiality and data privacy, participation and withdrawal details, contact information, consent declaration, and signature and date fields.
[DOCX File , 32 KB - aging_v8i1e63348_app1.docx]

Multimedia Appendix 2
Inclusion and exclusion criteria for trial participation.
[DOCX File , 28 KB - aging_v8i1e63348_app2.docx]

Multimedia Appendix 3

Think-aloud protocol.
[DOCX File , 30 KB - aging_v8i1e63348_app3.docx]

Multimedia Appendix 4
Interview guide for the focus groups.
[DOCX File , 32 KB - aging_v8i1e63348_app4.docx]

Multimedia Appendix 5
Complete acceptability results per category.
[DOCX File , 361 KB - aging_v8i1e63348_app5.docx]

Multimedia Appendix 6
Visual representations of the Microsoft Power BI (Microsoft Corp) analytics.
[DOCX File , 672 KB - aging_v8i1e63348_app6.docx]

Multimedia Appendix 7
Areas of improvement.
[DOCX File , 52 KB - aging_v8i1e63348_app7.docx]

Multimedia Appendix 8
Strengths, weaknesses, opportunities, and threats matrix of the More In Action app according to the Nonadoption, Abandonment, Scale-up, Spread, and Sustainability components.
[DOCX File , 31 KB - aging_v8i1e63348_app8.docx]

Multimedia Appendix 9
Willingness-to-pay analysis.
[DOCX File , 33 KB - aging_v8i1e63348_app9.docx]

References

1. Ageing. World Health Organization. URL: https://www.who.int/health-topics/ageing#tab=tab_1 [accessed 2023-12-01]
2. Bangsbo J, Blackwell J, Boraxbekk CJ, Caserotti P, Dela F, Evans AB, et al. Copenhagen Consensus statement 2019: physical activity and ageing. Br J Sports Med 2019 Jul 21;53(14):856-858 [FREE Full text] [doi: 10.1136/bjsports-2018-100451] [Medline: 30792257]
3. WHO guidelines on physical activity and sedentary behaviour. World Health Organization. 2020 Nov 25. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2023-12-01]
4. Ekelund U, Steene-Johannessen J, Brown WJ, Fagerland MW, Owen N, Powell KE, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. The Lancet 2016 Sep;388(10051):1302-1310. [doi: 10.1016/s0140-6736(16)30370-1]
5. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. The Lancet Global Health 2018 Oct;6(10):e1077-e1086. [doi: 10.1016/s2214-109x(18)30357-7]
6. Harvey JA, Chastin SF, Skelton DA. How sedentary are older people? A systematic review of the amount of sedentary behavior. J Aging Phys Act 2015 Jul;23(3):471-487. [doi: 10.1123/japa.2014-0164] [Medline: 25387160]
7. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424 [FREE Full text] [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
8. Bonnechère B, Kossi O, Mapinduzi J, Panda J, Rintala A, Guidetti S, et al. Mobile health solutions: an opportunity for rehabilitation in low- and middle income countries? Front Public Health 2022 Jan 24;10:1072322 [FREE Full text] [doi: 10.3389/fpubh.2022.1072322] [Medline: 36761328]
9. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
10. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 05;16 Suppl 5(Suppl 5):290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
11. Hsieh KL, Chen L, Sosnoff JJ. Mobile technology for falls prevention in older adults. J Gerontol A Biol Sci Med Sci 2023 May 11;78(5):861-868 [FREE Full text] [doi: 10.1093/gerona/glac116] [Medline: 35640254]

12. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. J Med Internet Res 2019 Nov 28;21(11):e14343 [FREE Full text] [doi: 10.2196/14343] [Medline: 31778121]
13. Sohaib Aslam A, van Luenen S, Aslam S, van Bodegom D, Chavannes NH. A systematic review on the use of mHealth to increase physical activity in older people. Clin eHealth 2020;3:31-39. [doi: 10.1016/j.CEH.2020.04.002]
14. Mönninghoff A, Kramer JN, Hess AJ, Ismailova K, Teepe GW, Tudor Car L, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699 [FREE Full text] [doi: 10.2196/26699] [Medline: 33811021]
15. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 06;3(1):250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
16. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
17. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User engagement and attrition in an app-based physical activity intervention: secondary analysis of a randomized controlled trial. J Med Internet Res 2019 Nov 27;21(11):e14645 [FREE Full text] [doi: 10.2196/14645] [Medline: 31774402]
18. Madeira RN, Germano H, Macedo P, Correia N. Personalising the user experience of a mobile health application towards patient engagement. Procedia Comput Sci 2018;141:428-433. [doi: 10.1016/j.procs.2018.10.173]
19. Daniels K, Lemmens R, Knippenberg E, Marinus N, Vonck S, Baerts J, et al. Promoting physical activity and a healthy active lifestyle in community-dwelling older adults: a design thinking approach for the development of a mobile health application. Front Public Health 2023 Nov 29;11:1280941 [FREE Full text] [doi: 10.3389/fpubh.2023.1280941] [Medline: 38106904]
20. Janols R, Sandlund M, Lindgren H, Pettersson B. Older adults as designers of behavior change strategies to increase physical activity-report of a participatory design process. Front Public Health 2022 Dec 21;10:988470 [FREE Full text] [doi: 10.3389/fpubh.2022.988470] [Medline: 36620266]
21. Nyman SR, Victor CR. Older people's participation in and engagement with falls prevention interventions in community settings: an augment to the Cochrane systematic review. Age Ageing 2012 Jan;41(1):16-23. [doi: 10.1093/ageing/afr103] [Medline: 21875865]
22. Lindsay S, Jackson D, Schofield G, Olivier P. Engaging older people using participatory design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX. [doi: 10.1145/2207676.2208570]
23. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
24. Mehra S, Visser B, Dadema T, van den Helder J, Engelbert RH, Weijs PJ, et al. Translating behavior change principles into a blended exercise intervention for older adults: design study. JMIR Res Protoc 2018 May 02;7(5):e117 [FREE Full text] [doi: 10.2196/resprot.9244] [Medline: 29720358]
25. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
26. Feeley N, Cossette S, Côté J, Héon M, Stremler R, Martorella G, et al. The importance of piloting an RCT intervention. Can J Nurs Res 2009 Jun;41(2):85-99. [Medline: 19650515]
27. Beyan N. The music methodology for usability measurement. In: Proceedings of the Posters and Short Talks of the 1992 SIGCHI Conference on Human Factors in Computing Systems. 1992 Presented at: CHI '92; May 3-7, 1992; Monterey, CA. [doi: 10.1145/1125021.1125114]
28. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
29. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Surrey, UK: Silverback Publishing; 2014.
30. Boen F, Seghers J, van Uffelen J. Motiveren tot verandering. In: Cahn W, Deenik J, Vermeulen J, editors. Leefstijlpsychiatrie. Houten, The Netherlands: Bohn Stafleu van Loghum; 2022.
31. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012 Jun 22;9(1):78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
32. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug 20;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]

33. Deci EL, Ryan RM. Motivation, personality, and development within embedded social contexts: an overview of self-determination theory. In: Ryan RM, editor. The Oxford Handbook of Human Motivation. Oxford, UK: Oxford University Press; 2012.
34. Aranha M, James K, Deasy C, Heavin C. Exploring the barriers and facilitators which influence mHealth adoption among older adults: a literature review. Gerontechnology 2021 Jan 01;20(2):1-16. [doi: 10.4017/gt.2021.20.2.424.06]
35. Delbaere K, Valenzuela T, Lord SR, Clemson L, Zijlstra GA, Close JC, et al. E-health StandingTall balance exercise for fall prevention in older people: results of a two year randomised controlled trial. BMJ 2021 Apr 06;373:n740 [FREE Full text] [doi: 10.1136/bmj.n740] [Medline: 33824131]
36. Muellmann S, Bragina I, Voelcker-Rehage C, Rost E, Lippke S, Meyer J, et al. Development and evaluation of two web-based interventions for the promotion of physical activity in older adults: study protocol for a community-based controlled intervention trial. BMC Public Health 2017 May 25;17(1):512 [FREE Full text] [doi: 10.1186/s12889-017-4446-x] [Medline: 28545506]
37. Brew-Sam N, Chib A, Torres AY, Ng JX, Wong YT, Sze-G Y. An integrated mHealth campaign to reduce the risk of falling for older adults. J Appl Gerontol 2022 May 27;41(5):1336-1347. [doi: 10.1177/07334648211062877] [Medline: 35085043]
38. Thompson C, Porter Starr KN, Kemp EC, Chan J, Jackson E, Phun J. Feasibility of virtually delivering functional fitness assessments and a fitness training program in community-dwelling older adults. Int J Environ Res Public Health 2023 May 30;20(11):5996 [FREE Full text] [doi: 10.3390/ijerph20115996] [Medline: 37297600]
39. Jungreitmayr S, Kranzinger C, Venek V, Ring-Dimitriou S. Effects of an app-based physical exercise program on selected parameters of physical fitness of females in retirement: a randomized controlled trial. Front Physiol 2022 Mar 4;13:821773 [FREE Full text] [doi: 10.3389/fphys.2022.821773] [Medline: 35317213]
40. Daly RM, Gianoudis J, Hall T, Mundell NL, Maddison R. Feasibility, usability, and enjoyment of a home-based exercise program delivered via an exercise app for musculoskeletal health in community-dwelling older adults: short-term prospective pilot study. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21094 [FREE Full text] [doi: 10.2196/21094] [Medline: 33439147]
41. van Dyck D, Herman K, Poppe L, Crombez G, de Bourdeaudhuij I, Gheysen F. Results of MyPlan 2.0 on physical activity in older Belgian adults: randomized controlled trial. J Med Internet Res 2019 Oct 07;21(10):e13219 [FREE Full text] [doi: 10.2196/13219] [Medline: 31593541]
42. Shake MC, Crandall KJ, Mathews RP, Falls DG, Dispennette AK. Efficacy of Bingocize: a game-centered mobile application to improve physical and cognitive performance in older adults. Games Health J 2018 Aug;7(4):253-261. [doi: 10.1089/g4h.2017.0139] [Medline: 30089018]
43. Konstantinidis EI, Billis AS, Mouzakidis CA, Zilidou VI, Antoniou PE, Bamidis PD. Design, implementation, and wide pilot deployment of FitForAll: an easy to use exergaming platform improving physical fitness and life quality of senior citizens. IEEE J Biomed Health Inform 2016 Jan;20(1):189-200. [doi: 10.1109/jbhi.2014.2378814]
44. van Het Reve E, Silveira P, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: part 2 of a phase II preclinical exploratory trial. J Med Internet Res 2014 Jun 25;16(6):e159 [FREE Full text] [doi: 10.2196/jmir.3055] [Medline: 24966165]
45. Bao T, Carender WJ, Kinnaird C, Barone VJ, Peethambaran G, Whitney SL, et al. Effects of long-term balance training with vibrotactile sensory augmentation among community-dwelling healthy older adults: a randomized preliminary study. J Neuroeng Rehabil 2018 Jan 18;15(1):5 [FREE Full text] [doi: 10.1186/s12984-017-0339-6] [Medline: 29347946]
46. Albergoni A, Hettinga FJ, Stut W, Sartor F. Factors influencing walking and exercise adherence in healthy older adults using monitoring and interfacing technology: preliminary evidence. Int J Environ Res Public Health 2020 Aug 24;17(17):6142 [FREE Full text] [doi: 10.3390/ijerph17176142] [Medline: 32846988]
47. Nikitina S, Didino D, Baez M, Casati F. Feasibility of virtual tablet-based group exercise among older adults in Siberia: findings from two pilot trials. JMIR Mhealth Uhealth 2018 Feb 27;6(2):e40. [doi: 10.2196/mhealth.7531] [Medline: 29487045]
48. Johnson N, Bradley A, Klawitter L, Johnson J, Johnson L, Tomkinson GR, et al. The impact of a telehealth intervention on activity profiles in older adults during the COVID-19 pandemic: a pilot study. Geriatrics (Basel) 2021 Jun 30;6(3):68 [FREE Full text] [doi: 10.3390/geriatrics6030068] [Medline: 34209416]
49. Papi E, Chiou SY, McGregor AH. Feasibility and acceptability study on the use of a smartphone application to facilitate balance training in the ageing population. BMJ Open 2020 Dec 02;10(12):e039054 [FREE Full text] [doi: 10.1136/bmjopen-2020-039054] [Medline: 33268409]
50. Taraldsen K, Mikolaizak AS, Maier AB, Mellone S, Boulton E, Aminian K, et al. Digital technology to deliver a lifestyle-integrated exercise intervention in young seniors-the PreventIT feasibility randomized controlled trial. Front Digit Health 2020 Jul 31;2:10 [FREE Full text] [doi: 10.3389/fdgth.2020.00010] [Medline: 34713023]
51. Lugade V, Torbitt M, O'Brien SR, Silsupadol P. Smartphone- and paper-based delivery of balance intervention for older adults are equally effective, enjoyable, and of high fidelity: a randomized controlled trial. Sensors (Basel) 2023 Aug 27;23(17):7451 [FREE Full text] [doi: 10.3390/s23177451] [Medline: 37687907]

52. Bosco A, McGarrigle L, Skelton DA, Laventure RM, Townley B, Todd C. Make movement your mission: evaluation of an online digital health initiative to increase physical activity in older people during the COVID-19 pandemic. Digit Health 2022;8:20552076221084468 [FREE Full text] [doi: 10.1177/20552076221084468] [Medline: 35295764]
53. Recio-Rodríguez JI, Gonzalez-Sanchez S, Tamayo-Morales O, Gómez-Marcos MA, Garcia-Ortiz L, Niño-Martín V, et al. Changes in lifestyles, cognitive impairment, quality of life and activity day living after combined use of smartphone and smartband technology: a randomized clinical trial (EVIDENT-Age study). BMC Geriatr 2022 Oct 06;22(1):782 [FREE Full text] [doi: 10.1186/s12877-022-03487-5] [Medline: 36203135]
54. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8(1):115 [FREE Full text] [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
55. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: two validation studies. J Sport Exerc Psychol 1991;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
56. Scherrenberg M, Falter M, Kaihara T, Xu L, van Leunen M, Kemps H, et al. Development and internal validation of the digital health readiness questionnaire: prospective single-center survey study. J Med Internet Res 2023 Mar 10;25:e41615 [FREE Full text] [doi: 10.2196/41615] [Medline: 36897627]
57. van Holle V, de Bourdeaudhuij I, Deforche B, van Cauwenberg J, van Dyck D. Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). BMC Public Health 2015 Apr 28;15(1):433 [FREE Full text] [doi: 10.1186/s12889-015-1785-3] [Medline: 25928561]
58. Meh K, Jurak G, Sorić M, Rocha P, Sember V. Validity and reliability of IPAQ-SF and GPAQ for assessing sedentary behaviour in adults in the European Union: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Apr 26;18(9):4602. [doi: 10.3390/ijerph18094602] [Medline: 33926123]
59. Chung PK, Leung KM. Psychometric properties of eight-item physical activity enjoyment scale in a Chinese population. J Aging Phys Act 2018 Oct 12;27(1):1-6. [doi: 10.1123/japa.2017-0212] [Medline: 29722637]
60. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
61. Fisher NI, Kordupleski RE. Good and bad market research: a critical review of Net Promoter Score. Appl Stochastic Models Bus Ind 2018 Nov 22;35(1):138-151. [doi: 10.1002/asmb.2417]
62. Baquero A. Net promoter score (NPS) and customer satisfaction: relationship and efficient management. Sustainability 2022 Feb 10;14(4):2011. [doi: 10.3390/su14042011]
63. Ensink CJ, Keijsers NL, Groen BE. Translation and validation of the System Usability Scale to a Dutch version: D-SUS. Disabil Rehabil 2024 Jan 27;46(2):395-400. [doi: 10.1080/09638288.2022.2160837] [Medline: 36573399]
64. Town P, Thabtah F. Data analytics tools: a user perspective. J Inf Knowl Manag 2019 Mar 22;18(01):1950002. [doi: 10.1142/s0219649219500023]
65. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3173618]
66. Jääskeläinen R. Think-aloud protocol. In: Handbook of Translation Studies. Amsterdam, The Netherlands: John Benjamins Publishing Company; 2010.
67. Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? Qual Res Sport Exerc Health 2017 May 26;9(4):514-531. [doi: 10.1080/2159676x.2017.1331501]
68. McDonald S, Edwards HM, Zhao T. Exploring think-alouds in usability testing: an international survey. IEEE Trans Prof Commun 2012 Mar;55(1):2-19. [doi: 10.1109/tpc.2011.2182569]
69. Ericsson KA. Valid and non-reactive verbalization of thoughts during performance of tasks: towards a solution to the central problems of introspection as a source of scientific data. J Conscious Stud 2003;10(9-10):1-18 [FREE Full text]
70. Dhakal K. NVivo. J Med Libr Assoc 2022 Apr 01;110(2):270-272 [FREE Full text] [doi: 10.5195/jmla.2022.1271] [Medline: 35440911]
71. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
72. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inf 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
73. Pruchno R. Technology and aging: an evolving partnership. Gerontologist 2019 Jan 09;59(1):1-5. [doi: 10.1093/geront/gny153] [Medline: 30629258]
74. Lee CC, Czaja SJ, Moxley JH, Sharit J, Boot WR, Charness N, et al. Attitudes toward computers across adulthood from 1994 to 2013. Gerontologist 2019 Jan 09;59(1):22-33 [FREE Full text] [doi: 10.1093/geront/gny081] [Medline: 29982458]
75. Nebeker C, Zlatar ZZ. Learning from older adults to promote independent physical activity using mobile health (mHealth). Front Public Health 2021 Aug 12;9:703910 [FREE Full text] [doi: 10.3389/fpubh.2021.703910] [Medline: 34476229]

76. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
77. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020 Sep 15;60(7):e513-e523 [FREE Full text] [doi: 10.1093/geront/gnz163] [Medline: 31773145]
78. Frennert S, Östlund B. Review: seven matters of concern of social robots and older people. Int J Soc Robot 2014 Jan 29;6(2):299-310. [doi: 10.1007/s12369-013-0225-8]
79. Degroote L, van Dyck D, de Bourdeaudhuij I, de Paepe A, Crombez G. Acceptability and feasibility of the mHealth intervention 'MyDayPlan' to increase physical activity in a general adult population. BMC Public Health 2020 Jun 29;20(1):1032 [FREE Full text] [doi: 10.1186/s12889-020-09148-9] [Medline: 32600352]
80. McGarrigle L, Todd C. Promotion of physical activity in older people using mHealth and eHealth technologies: rapid review of reviews. J Med Internet Res 2020 Dec 29;22(12):e22201 [FREE Full text] [doi: 10.2196/22201] [Medline: 33372894]
81. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. EHealth interventions for the promotion of physical activity in older adults: a systematic review. In: Proceedings of the 9th European Public Health Conference. 2016 Presented at: EPH 2016; November 9-12, 2016; Vienna, Austria. [doi: 10.1093/eurpub/ckw165.023]
82. Vasquez BA, Betriana F, Nemenzo E, Inabangan AK, Tanioka R, Garcia L, et al. Effects of healthcare technologies on the promotion of physical activities in older persons: a systematic review. Inform Health Soc Care 2023 Apr 03;48(2):196-210. [doi: 10.1080/17538157.2022.2086874] [Medline: 35699246]
83. Maes I, Mertens L, Poppe L, Crombez G, Vetrovsky T, van Dyck D. The variability of emotions, physical complaints, intention, and self-efficacy: an ecological momentary assessment study in older adults. PeerJ 2022;10:e13234 [FREE Full text] [doi: 10.7717/peerj.13234] [Medline: 35611175]
84. Maes I, Mertens L, Poppe L, Vetrovsky T, Crombez G, de Backere F, et al. Within-person associations of accelerometer-assessed physical activity with time-varying determinants in older adults: time-based ecological momentary assessment study. JMIR Aging 2023 Nov 23;6:e44425 [FREE Full text] [doi: 10.2196/44425] [Medline: 37995131]
85. Hernández-Reyes A, Molina-Recio G, Molina-Luque R, Romero-Saldaña M, Cámara-Martos F, Moreno-Rojas R. Effectiveness of PUSH notifications from a mobile app for improving the body composition of overweight or obese women: a protocol of a three-armed randomized controlled trial. BMC Med Inform Decis Mak 2020 Feb 24;20(1):40 [FREE Full text] [doi: 10.1186/s12911-020-1058-7] [Medline: 32093701]
86. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov 29;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]
87. Mair JL, Hayes LD, Campbell AK, Buchan DS, Easton C, Sculthorpe N. A personalized smartphone-delivered just-in-time adaptive intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. JMIR Form Res 2022 Apr 07;6(4):e34662 [FREE Full text] [doi: 10.2196/34662] [Medline: 35389348]
88. Rothman AJ, Sheeran P. The operating conditions framework: integrating mechanisms and moderators in health behavior interventions. Health Psychol 2021 Dec;40(12):845-857. [doi: 10.1037/hea0001026] [Medline: 32914997]
89. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr 01;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
90. Pickering TA, Huh J, Intille S, Liao Y, Pentz MA, Dunton GF. Physical activity and variation in momentary behavioral cognitions: an ecological momentary assessment study. J Phys Act Health 2016 Mar;13(3):344-351. [doi: 10.1123/jpah.2014-0547] [Medline: 26284314]
91. Dunton GF. Ecological momentary assessment in physical activity research. Exerc Sport Sci Rev 2017 Jan;45(1):48-54 [FREE Full text] [doi: 10.1249/JES.0000000000000092] [Medline: 27741022]
92. Firth J, Torous J, Yung AR. Ecological momentary assessment and beyond: the rising interest in e-mental health research. J Psychiatr Res 2016 Sep;80:3-4 [FREE Full text] [doi: 10.1016/j.jpsychires.2016.05.002] [Medline: 27236099]
93. de Marez L, Sevenhant R, Denecker F, Georges A, Wuyts G, Schuurman D. imec.digimeter 2023 : digitale trends in Vlaanderen. Imec. URL: https://biblio.ugent.be/publication/01HJ69C9YQEE4C4KXKBC7AYS2J [accessed 2023-12-01]
94. Grates MG, Heming AC, Vukoman M, Schabsky P, Sorgalla J. New perspectives on user participation in technology design processes: an interdisciplinary approach. Gerontologist 2019 Jan 09;59(1):45-57. [doi: 10.1093/geront/gny112] [Medline: 30260407]
95. Agan J, Koch LC, Rumrill PDJ. The use of focus groups in rehabilitation research. Work 2008 Jan 01;31(2):259-269. [doi: 10.3233/wor-2008-00731]
96. O.Nyumba T, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: insights from two decades of application in conservation. Methods Ecol Evol 2018 Jan 11;9(1):20-32. [doi: 10.1111/2041-210x.12860]
97. El-Kotob R, Giangregorio LM. Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. Pilot Feasibility Stud 2018 Aug 14;4(1):137 [FREE Full text] [doi: 10.1186/s40814-018-0326-0] [Medline: 30123527]

98. Shanyinde M, Pickering RM, Weatherall M. Questions asked and answered in pilot and feasibility randomized controlled trials. BMC Med Res Methodol 2011 Aug 16;11(1):117 [FREE Full text] [doi: 10.1186/1471-2288-11-117] [Medline: 21846349]
99. Bonnechère B, Timmermans A, Michiels S. Current technology developments can improve the quality of research and level of evidence for rehabilitation interventions: a narrative review. Sensors (Basel) 2023 Jan 12;23(2):875 [FREE Full text] [doi: 10.3390/s23020875] [Medline: 36679672]
100. Beukenhorst AL, Howells K, Cook L, McBeth J, O'Neill TW, Parkes MJ, et al. Engagement and participant experiences with consumer smartwatches for health research: longitudinal, observational feasibility study. JMIR Mhealth Uhealth 2020 Jan 29;8(1):e14368 [FREE Full text] [doi: 10.2196/14368] [Medline: 32012078]
101. Ceylana HH, Koseb B, Aydin M. Value based pricing: a research on service sector using Van Westendorp Price Sensitivity Scale. Procedia Soc Behav Sci 2014 Aug;148:1-6. [doi: 10.1016/j.sbspro.2014.07.013]
102. Cao W, Milks MW, Liu X, Gregory ME, Addison D, Zhang P, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. JMIR Mhealth Uhealth 2022 Mar 02;10(3):e29415 [FREE Full text] [doi: 10.2196/29415] [Medline: 35234655]
103. Nimmanterdwong Z, Boonviriya S, Tangkijvanich P. Human-centered design of mobile health apps for older adults: systematic review and narrative synthesis. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e29512 [FREE Full text] [doi: 10.2196/29512] [Medline: 35029535]
104. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of just-in-time adaptive interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019 Apr 03;16(1):31 [FREE Full text] [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
105. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018 Mar 22;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
106. Wilmink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. JMIR Aging 2020 Sep 10;3(2):e19554 [FREE Full text] [doi: 10.2196/19554] [Medline: 32723711]
107. Burnham JP, Lu C, Yaeger LH, Bailey TC, Kollef MH. Using wearable technology to predict health outcomes: a literature review. J Am Med Inform Assoc 2018 Sep 01;25(9):1221-1227 [FREE Full text] [doi: 10.1093/jamia/ocy082] [Medline: 29982520]
108. Baig MM, Afifi S, GholamHosseini H, Mirza F. A systematic review of wearable sensors and IoT-based monitoring applications for older adults - a focus on ageing population and independent living. J Med Syst 2019 Jun 15;43(8):233. [doi: 10.1007/s10916-019-1365-7] [Medline: 31203472]
109. Pelssers J, Delecluse C, Opdenacker J, Kennis E, van Roie E, Boen F. "Every step counts!": effects of a structured walking intervention in a community-based senior organization. J Aging Phys Act 2013 Apr;21(2):167-185. [doi: 10.1123/japa.21.2.167] [Medline: 22898451]
110. Boulton ER, Horne M, Todd C. Involving older adults in developing physical activity interventions to promote engagement: a literature review. Popul Ageing 2019 Jun 24;13(3):325-345. [doi: 10.1007/s12062-019-09247-5]
111. Bonnechère B. Unlocking the black box? A comprehensive exploration of large language models in rehabilitation. Am J Phys Med Rehabil 2024 Jan 12;103(6):532-537. [doi: 10.1097/phm.0000000000002440]
112. Golas SB, Nikolova-Simons M, Palacholla R, Op den Buijs J, Garberg G, Orenstein A, et al. Predictive analytics and tailored interventions improve clinical outcomes in older adults: a randomized controlled trial. NPJ Digit Med 2021 Jun 10;4(1):97 [FREE Full text] [doi: 10.1038/s41746-021-00463-y] [Medline: 34112921]
113. Nikolova-Simons M, Golas SB, den Buijs JO, Palacholla RS, Garberg G, Orenstein A, et al. A randomized trial examining the effect of predictive analytics and tailored interventions on the cost of care. NPJ Digit Med 2021 Jun 03;4(1):92 [FREE Full text] [doi: 10.1038/s41746-021-00449-w] [Medline: 34083743]
114. Chiao CY, Wu HS, Hsiao CY. Caregiver burden for informal caregivers of patients with dementia: a systematic review. Int Nurs Rev 2015 Sep 08;62(3):340-350. [doi: 10.1111/inr.12194] [Medline: 26058542]
115. Alltag S, Conrad I, Riedel-Heller SG. [Caregiver burden among older informal caregivers of patients with dementia and its influence on quality of life: a systematic literature review]. Z Gerontol Geriatr 2019 Aug;52(5):477-486. [doi: 10.1007/s00391-018-1424-6] [Medline: 29992357]
116. Yang Y, Koenigstorfer J. Determinants of physical activity maintenance during the Covid-19 pandemic: a focus on fitness apps. Transl Behav Med 2020 Oct 08;10(4):835-842 [FREE Full text] [doi: 10.1093/tbm/ibaa086] [Medline: 32926160]
117. Symons M, Meira Cunha C, Poels K, Vandebosch H, Dens N, Alida Cutello C. Physical activity during the first lockdown of the COVID-19 pandemic: investigating the reliance on digital technologies, perceived benefits, barriers and the impact of affect. Int J Environ Res Public Health 2021 May 22;18(11):5555 [FREE Full text] [doi: 10.3390/ijerph18115555] [Medline: 34067390]
118. Petersen JM, Kemps E, Lewis LK, Prichard I. Promoting physical activity during the COVID-19 lockdown in Australia: the roles of psychological predictors and commercial physical activity apps. Psychol Sport Exerc 2021 Sep;56:102002 [FREE Full text] [doi: 10.1016/j.psychsport.2021.102002] [Medline: 36567740]

119. Banskota S, Healy M, Goldberg E. 15 smartphone apps for older adults to use while in isolation during the COVID-19 pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525 [FREE Full text] [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
120. Ghani Z, Jarl J, Sanmartin Berglund J, Andersson M, Anderberg P. The cost-effectiveness of mobile health (mHealth) interventions for older adults: systematic review. Int J Environ Res Public Health 2020 Jul 22;17(15):5290 [FREE Full text] [doi: 10.3390/ijerph17155290] [Medline: 32708016]
121. Clet E, Leblanc P, Alla F, Cohidon C. Factors for the integration of prevention in primary care: an overview of reviews. BJGP Open 2024 Apr 05;8(3):BJGPO.2023.0141. [doi: 10.3399/bjgpo.2023.0141]

Abbreviations

AI: artificial intelligence
BCW: Behavior Change Wheel
CONSORT: Consolidated Standards of Reporting Trials
CSAT: Customer Satisfaction Score
DHRQ: Digital Health Readiness Questionnaire
IPAQ-SF: International Physical Activity Questionnaire–Short Form
mHealth: mobile health
MIA: More In Action
NASSS: Nonadoption, Abandonment, Scale-up, Spread, and Sustainability
NPS: Net Promoter Score
PA: physical activity
PACES: Physical Activity Enjoyment Scale
SUS: System Usability Scale

Edited by J Wang; submitted 17.06.24; peer-reviewed by N Santos, T Sun, M Shi, I Hedderson; comments to author 02.10.24; revised version received 23.10.24; accepted 03.12.24; published 27.01.25.

Please cite as:
Daniels K, Vonck S, Robijns J, Quadflieg K, Bergs J, Spooren A, Hansen D, Bonnechère B
Exploring the Feasibility of a 5-Week mHealth Intervention to Enhance Physical Activity and an Active, Healthy Lifestyle in Community-Dwelling Older Adults: Mixed Methods Study
JMIR Aging 2025;8:e63348
URL: https://aging.jmir.org/2025/1/e63348
doi:10.2196/63348
PMID:

Original Paper

Adapting the Technology Acceptance Model to Examine the Use of Information Communication Technologies and Loneliness Among Low-Income, Older Asian Americans: Cross-Sectional Survey Analysis

Pauline DeLange Martinez[1], PhD; Daniel Tancredi[2], PhD; Misha Pavel[3], PhD; Lorena Garcia[4], DrPH; Heather M Young[1], PhD

[1]Betty Irene Moore School of Nursing, University of California, Davis, Sacramento, CA, United States

[2]Department of Pediatrics, University of California, Davis, Sacramento, CA, United States

[3]Khoury College of Computer Science, Northeastern University, Boston, MA, United States

[4]Department of Public Health Sciences, University of California, Davis, Sacramento, CA, United States

Corresponding Author:
Pauline DeLange Martinez, PhD
Betty Irene Moore School of Nursing
University of California, Davis
2750 48th St
Sacramento, CA, 95817
United States
Phone: 1 9164262862
Email: pdmartinez@ucdavis.edu

Abstract

Background: Loneliness is a significant issue among older Asian Americans, exacerbated by the COVID-19 pandemic. Older age, lower income, limited education, and immigrant status heighten loneliness risk. Information communication technologies (ICTs) have been associated with decreased loneliness among older adults. However, older Asian Americans are less likely to use ICTs, particularly if they are immigrants, have limited English proficiency, or are low income. The Technology Acceptance Model posits that perceived usefulness (PU), and perceived ease of use (PEOU) are key factors in predicting technology use.

Objective: This study aimed to examine associations between PU, PEOU, ICT use, and loneliness among low-income, older Asian Americans.

Methods: Cross-sectional survey data were gathered from predominately older Asian Americans in affordable senior housing (N=401). Using exploratory factor analysis and Horn parallel analysis, we examined 12 survey items to identify factors accounting for variance in ICT use. We deployed structural equation modeling to explore relationships among the latent factors and loneliness, adjusting for demographic and cognitive factors.

Results: Exploratory factor analysis and Horn parallel analysis revealed 3 factors that accounted for 56.48% (6.78/12) total variance. PEOU combined items from validated subscales of tech anxiety and comfort, accounting for a 28.44% (3.41/12) variance. ICT use combined years of technological experience, computer, tablet, and smartphone use frequency, accounting for 15.59% (1.87/12) variance. PU combined 2 items assessing the usefulness of technology for social connection and learning and accounted for a 12.44% (1.49/12) variance. The 3-factor structural equation modeling revealed reasonable fit indexes (χ^2_{133}=345.132; *P*<.001, chi-square minimum (CMIN)/df = 2595, comparative fit index (CFI)=0.93, Tucker-Lewis Index (TLI)=0.88). PEOU was positively associated with PU (β=.15; *P*=.01); PEOU and PU were positive predictors of ICT use (PEOU β=.26, *P*<.001; PU β=.18, *P*=.01); and ICT use was negatively associated with loneliness (β=–.28, *P*<.001). Demographic and health covariates also significantly influenced PU, PEOU, ICT use, and loneliness. English proficiency and education positively predicted PEOU (*r*=0.25, *P*<.001; *r*=0.26, *P*<.001) and ICT use (β=1.66, *P*=.03; β=.21, *P*<.001), while subjective cognitive decline and Asian ethnicity were positively associated with loneliness (β=.31, *P*<.001; β=.25, *P*<.001).

Conclusions: This study suggests that targeted interventions enhancing PU or PEOU could increase ICT acceptance and reduce loneliness among low-income Asian Americans. Findings also underscore the importance of considering limited English proficiency and subjective cognitive decline when designing interventions and in future research.

(JMIR Aging 2025;8:e63856) doi:10.2196/63856

KEYWORDS

social isolation; loneliness; aged; older adults; Asian American; immigrant; vulnerable populations; internet; information and communication technologies; ICTs; digital divide; technology acceptance model; mobile phone

Introduction

Background

The 2019 California Health Interview Survey (CHIS) found that 1 in 4 (25.7%) Asian Americans aged 65 years and older were lonely [1]. Loneliness is defined as a subjective experience stemming from perceived isolation or a disparity in one's desired and actual social interactions [2]. In the CHIS, Asian American older adults reported significantly lower levels of perceived social and emotional support (56%) as compared with non–Asian American older adults (80%) [3]. During the COVID-19 pandemic, loneliness increased among older Americans, and even today, loneliness levels are higher as compared with prepandemic levels [4]. In 2023, the Surgeon General announced that the United States is experiencing a pandemic of loneliness [2].

As compared with younger generations, older adults are particularly at risk for loneliness and social isolation due to factors such as retirement, relocation, and shrinking social circles. Besides older age, other risk factors for loneliness include financial insecurity, low educational attainment, poor physical or mental health, being an immigrant, having a disability, and living alone [2,5-7]. A systematic review exploring factors associated with loneliness among older Asian American immigrants found that migration grief, diminished ethnic ties, mental and physical impairment, deteriorating health conditions, living alone, a lack of meaningful social connections and support networks, and fewer interactions with family members were all significant factors contributing to loneliness [8].

Among older adults, the use of information communication technologies (ICTs), including smartphones, tablets, personal computers, the internet, and social media, is associated with decreased loneliness [9,10]. ICT use can strengthen preexisting connections with family and friends, foster new social relationships, and build intergenerational bonds [9]. ICT use is also positively associated with self-efficacy, self-esteem, a sense of autonomy, independence, and greater well-being among older adults [9]. Older Asian Americans can further benefit from using ICTs to stay in contact with distant relatives and maintain a connection with their culture of origin; this is particularly relevant given that 85% of Asian Americans aged 65 years and older are foreign-born [11-13]. Furthermore, ICTs can facilitate access to information in one's native language and translation services. However, both the CHIS and the National Health and Aging Trends Study (NHATS) showed that older Asian Americans are less likely than non-Hispanic White older adults to use the internet, send emails or text messages, conduct personal tasks on the internet, or seek web-based health information, particularly if they are immigrants, have limited English proficiency, or are low income [14-16]. Furthermore, other factors, such as age, gender, educational attainment, and subjective cognitive decline significantly impact ICT acceptance and use among low-income, older Asian Americans [17,18].

A recent systematic mapping review identified 59 articles describing 119 factors that predict older adults' intention to use digital technologies [19]. However, despite a rich literature focused on this topic, the mapping revealed that most studies (68%) did not examine these factors using an established theoretical framework or model for technology acceptance. For example, although loneliness was identified as a factor associated with ICT use, it has not been analyzed through a theoretical framework or model [19].

Theoretical Framework

The Technology Acceptance Model (TAM) is the most commonly used theoretical model to study technology adoption among older adults and also among the general population [19-21]. The backbone of the TAM is comprised of perceived usefulness (PU) and perceived ease of use (PEOU). PU relates to one's perception of technology as being useful for accomplishing desired goals, while PEOU refers to how much effort one anticipates needing to make to learn to use new technology. The TAM proposes that PEOU predicts PU, PU, and PEOU predict attitudes toward technology, and these attitudes predict behavioral intention to use technology, which subsequently influences actual use [21].

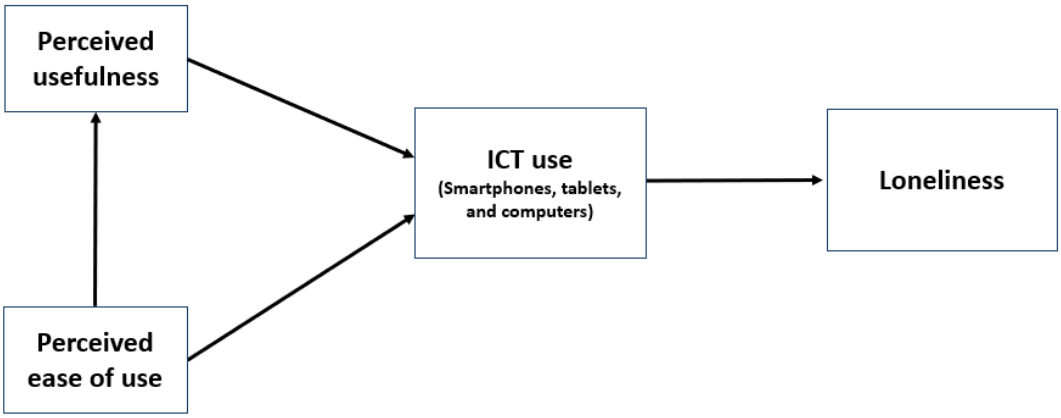
In 2 previous studies, we validated a simplified TAM to predict ICT use among low income, older Asian Americans. The simplified model removed the mediators (attitudes toward technology and behavioral intentions) and adapted the constructs of PU and PEOU from the original TAM [17,18]. In the adapted model, PU was defined as older adults' perceptions of ICTs as being useful for connecting with family and friends and learning new information and skills, and PEOU was operationalized using 6 evidence-based items that measure older adults' comfort and anxiety with ICTs [18]. However, in our previous work, we had not empirically examined the assumptions of the operationalization of PU, PEOU, and ICT use using robust statistical methods. In addition, the association between ICT use and loneliness among older Asian Americans has never previously been examined using the TAM framework.

Research Design

This cross-sectional study aimed to rigorously assess the assumptions underlying the conceptualization of PU, PEOU,

and ICT use using a series of statistical techniques. In addition, we plan to extend the simplified TAM as described in DeLange Martinez et al [18] to explore the association between technology acceptance and loneliness among low income, older Asian Americans (Figure 1). Based on previous studies, we adjusted the model for age, gender, education, English proficiency, Asian ethnicity, and subjective cognitive decline [17,18].

Figure 1. Extended Technology Acceptance Model examining the association between information communication technologies acceptance and loneliness among low income, older Asian Americans. ICT: information communication technology.



In this study, we test three hypotheses, that are (1) H1: PEOU will be positively associated with PU. (2) H2: PEOU and PU will be significant, positive predictors of ICT use. (3) H3: ICT use will be significantly negatively associated with loneliness.

Methods

Data and Sample

In the Fall of 2020, The Lighthouse Project for Older Adults was launched with the aim of developing a scalable model to address barriers to technology use among residents of affordable senior housing communities. To inform the intervention, focus groups were held with 29 residents and 13 staff across 2 communities. The discussion revealed many challenges, including high rates of social isolation, low literacy levels in multiple languages, cognitive and sensory challenges, low comfort with technology, and lack of infrastructure [22].

To address these challenges, residents were offered access to high-speed broadband internet, ICT devices, a series of digital literacy training courses, and tech support led by peer ambassadors. As part of the program evaluation, participants completed pre- and postsurveys. The surveys were evidence-based, translated into 5 languages (English, Spanish, Chinese, Vietnamese, and Korean), and pilot-tested with 20 residents and 4 staff. The final surveys were self-administered by all Lighthouse participants in their preferred language with staff available to assist as needed.

Measures

The operationalization of PU, PEOU, and ICT use were based on the theoretical constructs from the original TAM [21]. In 2 previous studies, we operationalized these 3 constructs with 12 survey items as described below, standardizing and summing the items for subsequent analysis with hierarchical linear regression [17,18]. In this study, we examined the theoretical assumptions of these constructs by using exploratory factor analysis (EFA), Horn parallel analysis, and structural equation modeling (SEM).

Perceived Usefulness

In the TAM, PU refers to whether one perceives technology to be useful for accomplishing desired goals [21]. In this study, we operationalized PU with 2 items developed by Sims et al [23] in their measure of Motivations for ICT Use. (“Technology helps me be connected with family and friends,” and “Technology helps me learn new information and skills.”) Response categories for both statements ranged on a scale of 1 (strongly disagree) to 4 (strongly agree).

Perceived Ease of Use

In the TAM, the construct of PEOU refers to how much effort one anticipates needing to make to learn to use a new technology. PEOU has been operationalized to measure feelings of confusion, frustration, or ease when using technology; predictability or intuitiveness of the system; and frequency of making mistakes [21]. In this study, we examined 6 items from 2 validated subscales that we predicted would collectively represent the construct of PEOU. Two items were included from the Senior TAM – Tech Anxiety Subscale [24] (“I feel apprehensive about using technology,” and, “I hesitate to use technology for fear of making mistakes I cannot correct.”) and 4 items were included from the Attitudes Towards Computers Questionnaire (ATCQ) – Comfort Subscale [25] (“I feel comfortable with technology” (reverse scored), “Technology makes me nervous,” “I don’t feel confident about my ability to use technology,” and, “Technology is confusing”). Response options for all statements ranged on a scale from 1 (strongly agree) to 4 (strongly disagree). Although the original ATCQ - Comfort Subscale included 5 items, one item was not included in the Lighthouse for Older Adults survey (“Computers make

me feel dumb.") based on input from staff about cultural relevance.

ICT Use

We operationalized ICT use with 4 survey items. Three items asked about the frequency of use of computers, tablets, and smartphones. ("How often do you use a desktop or laptop computer?"; "How often do you use a tablet or iPad?"; and "How often do you use a smartphone (iPhone or Android)?") Response options ranged from 0 (never, or I do not own), to 3 (about once per day). A fourth item inquired about years of experience using ICTs ("How long have you been using technology, such as a computer, laptop, tablet, or smartphone?"). Response options ranged from 0 (I've never used these) to 3 (more than 2 years).

Dependent Construct: Loneliness

Loneliness was measured using the 3-item University of California, Los Angeles (UCLA) Loneliness Scale ("How often do you feel that you lack companionship?"; "How often do you feel left out?"; and "How often do you feel isolated from others?") [26]. While the University of California, Los Angeles Loneliness Scale typically includes 3 response options (hardly ever, some of the time, often), a fourth answer option (never) was added because several residents handwrote "never" in the margins of the survey during pilot testing. During analysis, "never" responses were collapsed with "hardly ever." Therefore, in our analysis, response categories ranged on a scale of 1 (never or hardly ever) to 3 (often).

Control Variables

In our final SEM, we controlled for age, gender, Asian ethnicity, education, subjective cognitive decline (measured with 1 item, "During the past 12 months, have you experienced confusion or changes in memory that is happening more often or is getting worse?" [dichotomous response options]), and English proficiency (measured with one item, "How well do you speak English?" (response options ranged from 1 [not at all] to 4 [very well]). The full survey used in the Lighthouse Project for Older Adults is attached as a Multimedia Appendix 1.

Analytic Strategy

We began our analysis by examining descriptive statistics and conducting Pearson correlation analysis to explore relationships among all variables.

Subsequently, using IBM SPSS Statistics (version 29), we performed EFA to investigate the underlying theoretical constructs of the survey items concerning attitudes and use of ICTs. Despite having hypotheses about the latent variables, we chose to conduct EFA before confirmatory factor analysis (CFA), aiming for a more data-driven approach to thoroughly explore the underlying structure. To determine the number of factors to retain, we applied several criteria, including factors with loadings greater than .45 and using Kaiser eigenvalues-greater-than-one rule as illustrated on a scree plot [27]. We used principal axis factoring and varimax rotation with Kaiser normalization to account for the correlational nature of the factors. In addition, Horn parallel analysis was used to confirm the EFA results.

Moving forward, we used IBM SPSS AMOS (version 29) to conduct SEM, beginning with CFA. SEM integrates measurement models and structural models, allowing for validation of instruments and analysis of relationships while considering variances and covariances. It facilitates the examination of complex relationships among multiple variables and enables mediation analysis [28]. During CFA, we assessed convergent validity using criteria suggested by Hair et al [29] and Fornell and Larcker [30], calculating average variance extracted (AVE) and composite reliability. Discriminant validity was determined based on the Fornell-Larcker criterion [30].

Next, we used SEM to represent and test our 3 hypotheses, exploring relationships among PEOU, PU, ICT use, and our outcome of interest (loneliness). As shown in Figure 1, we examined a model with PU partially mediating the relationship between PEOU and ICT use, and ICT use fully mediating the relationships between PU, PEOU, and loneliness. We also examined the impact of adjusting for demographic and cognitive factors based on previous findings [15-18]. Maximum likelihood estimation was used for factor structure verification, and missing data were imputed by estimating means and intercepts. Measures of fit, including chi-square statistic (χ^2), chi-square divided by degrees of freedom (χ^2/df), comparative fit index (CFI), Tucker-Lewis coefficient (TLI), and root-mean-square error of approximation (RMSEA), were reported. Regression weights and correlation estimates among latent factors, the outcome of interest, and control factors were also provided. For all analyses, the α-level for testing significance was set to .05.

Ethical Considerations

For this study, we analyzed presurveys from 5 Lighthouse communities, collected between July 2021 and July 2022, before receiving the intervention. All participants were aged 62 years and older, based on housing eligibility criteria. In total, 31 participants were excluded from the analysis due to missing at least 1 of the dependent variables. The final dataset included 401 participants. On the basis of the HRP-210 Determination Request, the University of California, Davis, institutional review board determined that this research is exempt as it did not directly involve human participants and used deidentified secondary data (ID: 1938286-1).

Results

Overview

Participant (n=401) demographics are described in Table 1. Participants ranged in age from 62 to 97 (mean 79.07, SD 7) years, most were female, had a high school degree or less, reported limited English proficiency, and were Asian. Over a quarter of participants reported subjective cognitive decline and over a third reported loneliness.

Table 1. Participant demographics (N=401).

Participant demographics	Frequency, n	Percentage, %
Gender		
Female	276	69.2
Male	123	30.8
Ethnicity		
Non-Hispanic White	48	12.0
Asian	320	79.8
Korean	262	65.3
Chinese	50	12.5
Japanese	2	0.5
Filipino	2	0.5
Vietnamese	1	0.2
Latinx	11	2.7
Black or African American	6	1.5
American Indian or Alaskan Native	2	0.5
More than one race or ethnicity	7	1.7
English proficiency		
Very well	58	14.8
Well	62	15.8
Not well	172	43.8
Not at all	101	25.7
Educational attainment		
Never attended school	21	5.5
Some high school	127	33.0
Completed high school or general educational development	72	18.7
Some college	75	19.5
College degree	65	16.9
Graduate degree	25	6.5
Subjective cognitive decline		
No	283	72.2
Yes	109	27.8
Years of experience using information communication technologies		
More than 2 years	239	59.6
1 to 2 years	39	9.7
Less than 1 year	32	8.0
I have never used these	91	22.7
Computer use		
About once per day	87	21.7
2 to 4 times per week	28	7.0
Once or less than once per week	25	6.2
Never	261	65.1
Tablet use		
About once per day	92	22.9

Participant demographics	Frequency, n	Percentage, %
2 to 4 times per week	29	7.2
Once or less than once per week	25	6.2
Never	255	63.6
Smartphone use		
About once per day	247	61.6
2 to 4 times per week	39	9.7
Once or less than once per week	16	4.0
Never	99	24.7

About 1 in 5 participants reported they have never used a smartphone, tablet, or computer, while 6 in 10 reported over 2 years of experience using these devices. When asked about the frequency of use of each type of device, two-thirds reported that they never use a computer, slightly less than two-thirds never use a tablet, and a quarter never use a smartphone. Despite this, 89.2% (340/381) somewhat or strongly agreed that technology is useful for connecting with family and friends, and 90.4% (341/377) agreed that technology is useful for learning new information and skills.

Next, we used correlation analysis to examine relationships among items measuring ICT use, PEOU, PU, loneliness, and our control variables (age, gender, Asian ethnicity, education, subjective cognitive decline, and English proficiency). The assumptions for factor analysis were met. We observed multiple correlations among the items, most ranging from 0.30 and above. Importantly, none of these correlations exceeded 0.9, indicating no issues with multicollinearity. Correlation results are included in Multimedia Appendix 2.

EFA and Horn Parallel Analysis

The Kaiser-Meyer-Olkin measure of sampling adequacy was .83 indicating sufficient correlation among the variables. The Bartlett test of Sphericity indicated $P<.001$, allowing us to reject the null hypothesis that the correlation matrix is an identity matrix and that it is reasonable to proceed with EFA.

Three factors had an initial eigenvalue greater than one (Factor 1=4.65, Factor 2=1.89, and Factor 3=1.50). This is illustrated with the scree plot (Figure 2). The rotated factor matrix is shown in Table 2.

Figure 2. Scree plot.

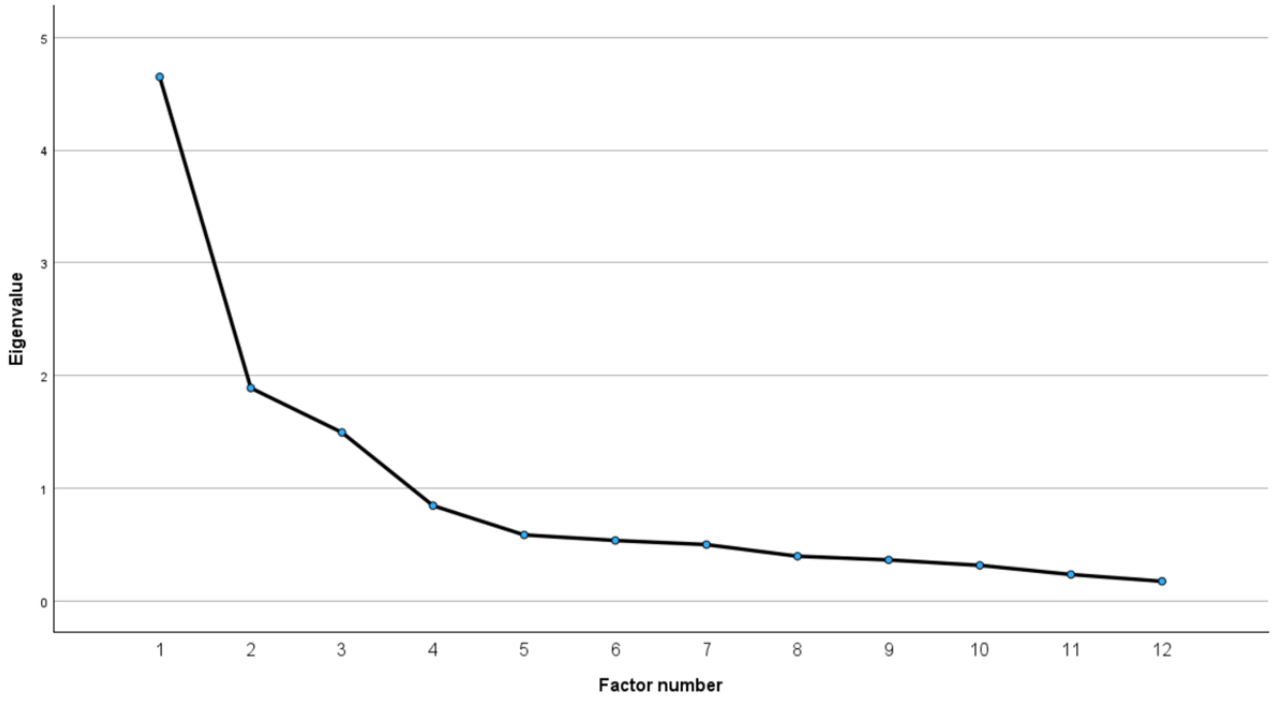


Table 2. Rotated factor matrix of independent variables.

Variables	Factor		
	PEOU[a]	ICT[b] use	PU[c]
Tech comfort	.40	.42	.35
Tech nervous	.72	.08	.01
Tech confidence	.82	.16	.10
Tech confusion	.88	.14	.08
Tech apprehension	.79	.20	.10
Tech fear	.75	.22	.10
Tech connection	.06	.07	.75
Tech learning	.09	.15	.85
Years of tech experience	.21	.60	.16
Computer use frequency	.19	.67	.03
Tablet use frequency	.03	.68	–.03
Smartphone use frequency	.11	.52	.17

[a]PEOU: perceived ease of use.

[b]ICT: Information communication technology.

[c]PU: perceived usefulness.

The first factor, PEOU, combined the 2 items from the Senior TAM – Tech Anxiety Subscale and 3 of the 4 items from the Attitudes Towards Computers Questionnaire – Comfort Subscale. One item, "I feel comfortable with technology," had a factor loading of .40. According to Tabachnick and Fidell [27], loading above 0.71 is excellent, 0.63 is very good, 0.55 is good, 0.45 fair, and 0.32 poor. Therefore, we dropped the tech comfort item from PEOU in further analyses.

The second factor, ICT use, combined 4 items: years of tech experience, computer use frequency, tablet use frequency, and smartphone use frequency.

The third factor, PU, combined 2 items: "Technology helps me be connected with family and friends," and "Technology helps me learn new information and skills."

After rotation, the 3 factors combined accounted for a total variance of 56.48% (6.78/12), with PEOU accounting for 28.44% (3.41/12) of the variance, ICT use for 15.59% (1.87/12), and PU for 12.44% (1.49/12). The 3 factors were confirmed when running a Horn parallel analysis. These factors also showed high internal reliability with Cronbach α scores of .90, .74, and .76 for PEOU, ICT use, and PU, respectively.

EFA was conducted separately for the dependent variables and generated 1 interpretable factor, loneliness, with an eigenvalue of 2.29 (Table 3). Loneliness had strong internal reliability with a Cronbach α score of .84.

Table 3. Component matrix for dependent variables using principal extraction method.

Variables	Factor (loneliness)
Lack companionship	.69
Feel left out	.93
Feel isolated	.79

CFA Findings

We ran CFA to further examine the relationships among the latent variables and to assess our conceptual model (Figure 3). As shown in Table 4, each of the items significantly loaded to form the 3 latent factors, confirming our EFA results.

Except for chi-square, all fit indices reached recommended level of fit: (χ^2_{41}=182.114; P<.001, chi-square minimum (CMIN)/df=4.44, CFI=0.92, TLI=0.87). The RMSEA of 0.09 was borderline. Since χ^2 is sensitive to large sample sizes, with a large sample of 401 participants, it was not unusual to get a significant value; for sample sizes greater than 250, a significant χ^2 value is acceptable [31].

Figure 3. Standardized results from confirmatory factor analysis to assess our conceptual model. ICT: information communication technology.

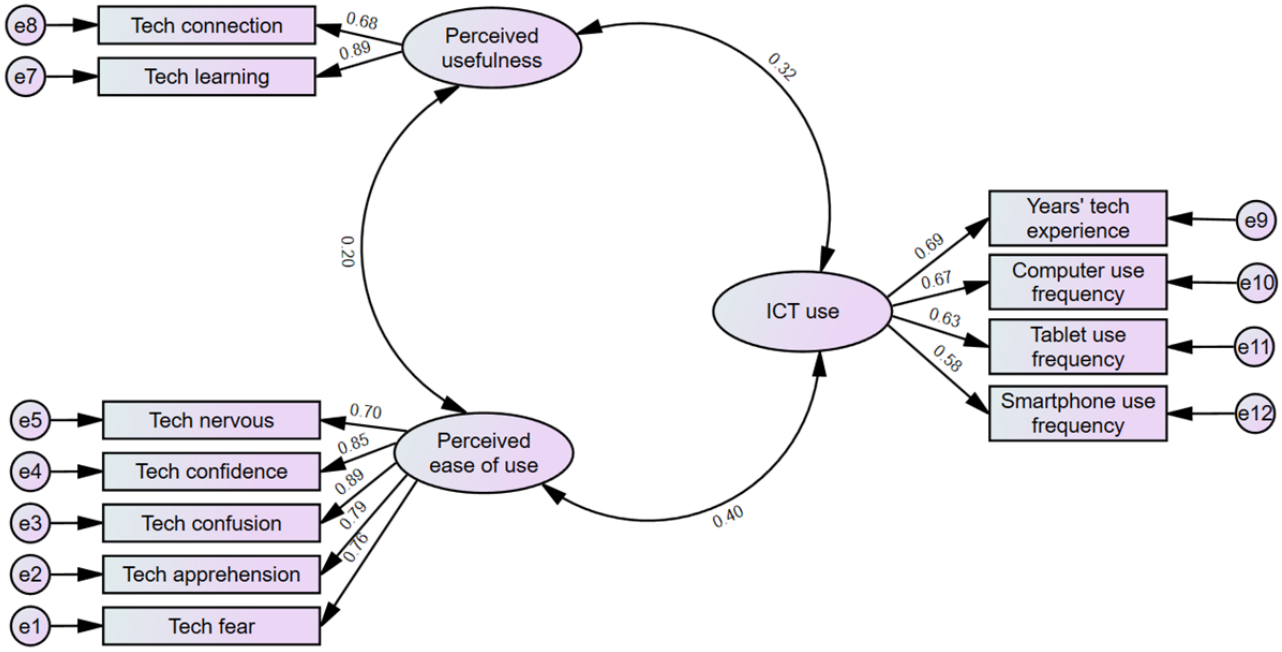


Table 4. Regression estimates of items loading into 3 factors.

Observed variables		Latent variables	Standardized regression estimate (factor loading)	Unstandardized regression estimate	SE	Composite reliability	*P* value
Tech fear	←	PEOU[a]	.76	1.00	—[b]	—	—
Tech apprehension	←	PEOU	.80	.99	.06	15.963	<.001
Tech confusion	←	PEOU	.89	1.09	.06	18.156	<.001
Tech confidence	←	PEOU	.85	1.05	.06	17.130	<.001
Tech nervous	←	PEOU	.70	.86	.06	13.797	<.001
Tech learning	←	PU[c]	.90	1.00	—	—	—
Tech connection	←	PU	.68	.77	.16	4.796	<.001
Years of tech experience	←	ICT[d] use	.69	1.00	—	—	—
Computer use frequency	←	ICT use	.67	.97	.10	10.053	<.001
Tablet use frequency	←	ICT use	.63	.93	.10	9.719	<.001
Smartphone use frequency	←	ICT use	.58	.86	.09	9.166	<.001

[a]PEOU: perceived ease of use.

[b]Not applicable.

[c]PU: perceived usefulness.

[d]ICT: Information communication technology.

There was evidence for convergent validity because all three of the conditions were fulfilled, that is, (1) composite reliability values are 0.7 or greater, (2) all standardized factor loadings are 0.5 or greater, and (3) AVE values are 0.5 or greater [29]. All 3 of these criteria were met for PEOU and PU, which had composite reliability values of .90 and .77, standardized factor loadings all greater or equal to .68, and AVE values of .64 and .63, respectively.

ICT use had an AVE of .42, slightly lower than ideal [29]. However, according to Fornell and Larcker [30], if the AVE is less than 0.5, but composite reliability is higher than 0.6, the convergent validity of the construct is acceptable [30]. ICT use had a composite reliability of .74, therefore we concluded that the latent variable had convergent validity.

Discriminant validity was met, with discriminant values of .80, .80, and .65 for PEOU, PU, and ICT use, respectively, while the correlation estimates all fell below 0.4 as shown in Table 5 [30].

Finally, as shown in Table 6, the relationships between the latent variables were all significant.

Table 5. Correlations and covariances of latent variables in confirmatory factor analysis.

Latent variables			Correlation estimate	Covariance estimate	SE	Composite reliability	*P* value
PEOU[a]	↔	PU	.200	.107	.033	3.273	.001
ICT[b] Use	↔	PU	.317	.190	.042	4.501	<.001
ICT Use	↔	PEOU	.399	.269	.048	5.643	<.001

[a]PEOU: perceived ease of use.

[b]ICT: Information communication technologies.

Table 6. Regression weights (unadjusted model).

Latent variables			Standardized regression estimate	Unstandardized regression estimate	SE	Composite reliability	*P* value
PU[a]	←	PEOU	.20	.18	.05	3.38	<.001
ICT[b] Use	←	PU	.24	.31	.10	3.19	.001
ICT Use	←	PEOU	.37	.41	.07	5.75	<.001
Loneliness	←	ICT Use	-.29	-.15	.03	-4.48	<.001

[a]PU: perceived usefulness.

[b]ICT: Information communication technologies.

Structural Equation Modeling

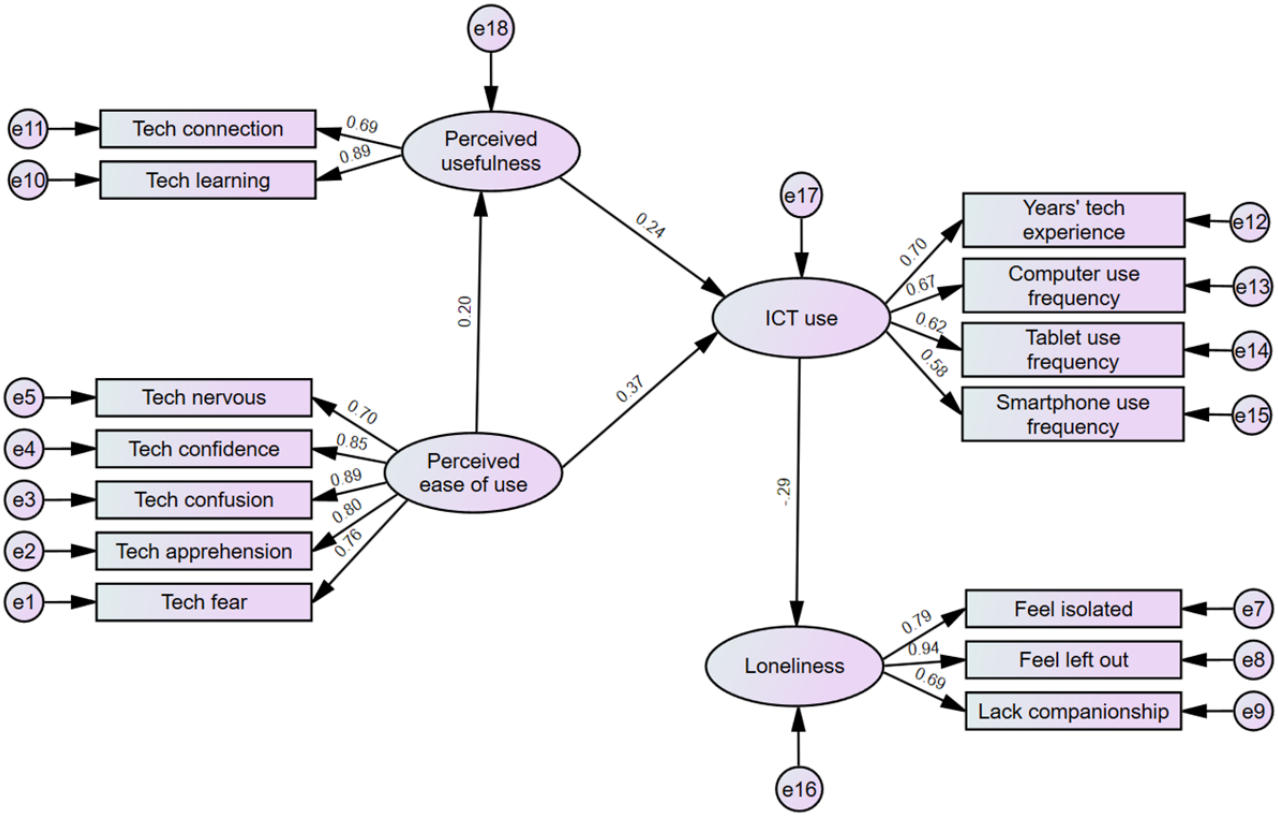
The 3-factor SEM revealed reasonable fit indexes (χ^2_{73}=231.835, CMIN/DF=3.18, CFI=0.93, TLI=0.90). An RMSEA of 0.07 was acceptable. Once again, the chi-square value was significant, which, as previously described, is acceptable for this sample size [31].

As shown in Figure 4 and Table 6, hypothesis 1 was supported; PEOU was significantly, and positively associated with PU (β=.20, *P*<.001).

Hypothesis 2 was also supported; both PEOU and PU were significant, positive predictors of ICT use (PEOU: β=.37; *P*<.001; PU: β=.24; *P*=.001).

Finally, hypothesis 3 was supported; ICT use was significantly, negatively associated with loneliness (β=–.29; *P*<.001).

Figure 4. Standardized results from structural equation modeling (unadjusted model).



Structural Equation Modeling Adjusting for Control Factors

When we reran the model adjusting for age, gender, education, Asian ethnicity, English proficiency, and subjective cognitive decline, the model fit improved (Figure 5; χ^2_{133}=345.13, *P*<.001, CMIN/DF=2595, CFI=0.93, TLI=0.88). The RMSEA of 0.06 was acceptable.

As shown in Table 7, the adjusted results continued to support Hypotheses 1, 2, and 3. PEOU continues to be significantly positively associated with PU (β=.152, *P*=.01). PEOU and PU were significant, positive predictors of ICT use (PEOU: β=.26, *P*<.001; PU: β=.179, *P*=.01). And, ICT use was significantly negatively associated with loneliness (β=–.28, *P*<.001). In addition, some of the control variables were significant predictors of the endogenous variables, PU and ICT use, and the dependent variable, loneliness (Table 7). Education was significantly, positively associated with PU (β=.19; *P*=.003). English proficiency and education significantly, positively predicted ICT use (English proficiency: β=1.66; *P*=.03; Education: β=.21; *P*<.001), while age was negatively associated with ICT use (β=–1.36; *P*=.01). Finally, subjective cognitive decline and Asian ethnicity were each positively associated with loneliness (subjective cognitive decline: β=.31; *P*<.001; Asian ethnicity: β=.25; *P*<.001).

Figure 5. Adjusted, standardized results from structural equation modeling. ICT: information communication technology.

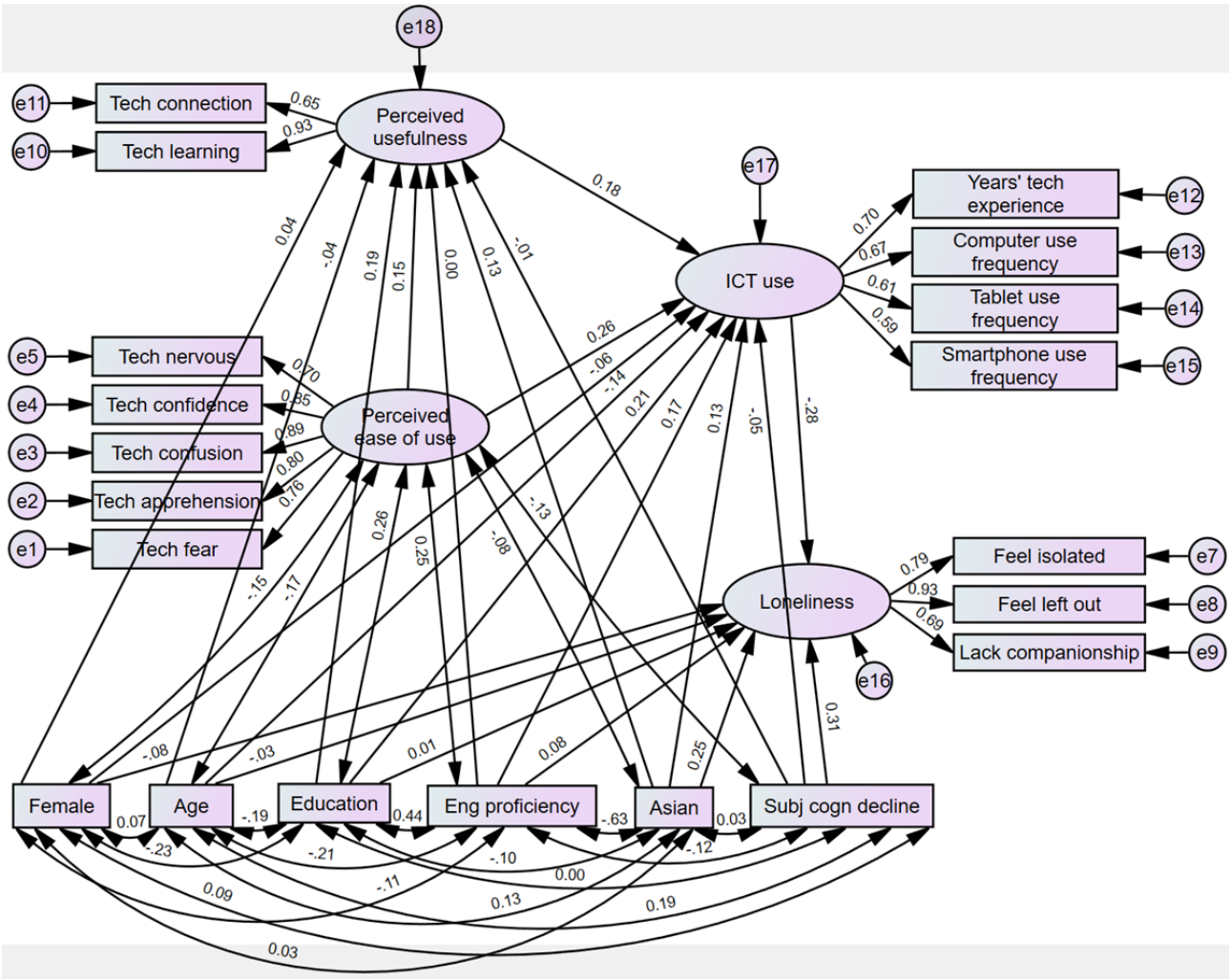


Table 7. Regression weights (adjusted model).

Variables	Standardized regression estimate	Unstandardized regression estimate	SE	Composite reliability	*P* value
PU[a]←PEOU[b]	.15[c]	.14	.06	2.55	.01
PU←English proficiency	–.003	–.002	.06	–.04	.97
PU←Asian	.13	.24	.13	1.83	.07
PU←subjective cognitive decline	–.01	–.02	.09	–.22	.83
PU←age	–.04	–.004	.01	–.78	.44
PU←education	.19[c]	.10	.03	2.98	.003
PU←female	.04	.07	.09	.80	.42
ICT[d] use←PU	.18[c]	.22	.08	2.70	.01
ICT use←female	–.06	–.12	.10	–1.17	.24
ICT use←age	–.14[c]	–.02	.01	–2.52	.01
ICT use←English proficiency	.17[c]	.15	.07	2.13	.03
ICT use←subjective cognitive decline	–.05	–.10	.10	–.92	.36
ICT use←PEOU	.26[c]	.29	.07	4.30	<.001
ICT use←education	.21[c]	.13	.04	3.30	<.001
ICT use←Asian	.13	.29	.15	1.93	.05
Loneliness←female	–.08	–.08	.05	–1.59	.11
Loneliness←age	–.03	–.002	.003	–.63	.53
Loneliness←English proficiency	.08	.04	.04	1.01	.31
Loneliness←ICT use	–.28[c]	–.14	.04	–3.93	<.001
Loneliness←subjective cognitive decline	.31[c]	.31	.05	5.76	<.001
Loneliness←education	.01	.004	.02	.21	.84
Loneliness←Asian	.25[c]	.28	.08	3.64	<.001

[a]PU: perceived usefulness.

[b]PEOU: perceived ease of use.

[c]$P \leq .05$.

[d]ICT: Information communication technology.

While PEOU was an exogenous factor, correlation estimates reveal significant associations with all the control variables except for Asian ethnicity (Table 8). English proficiency and education were each significantly positively associated with PEOU (English proficiency: r=0.25, P<.001; Education: r=0.26; P<.001). Subjective cognitive decline, age, and female gender were each significantly, and negatively associated with PEOU (Subjective cognitive decline: r=–0.130; P=.02; Age: r=–0.18; P=.001; Female: r=–0.15; P=.01).

Table 8. Correlations and covariances of control variables with perceived ease of use (adjusted model).

Control variables	Correlation estimate	Covariance estimate	SE	Composite reliability	*P* value
Asian	–.09	–.03	.02	–1.58	.11
Subjective cognitive decline	–.13	–.05	.02	–2.40	.02
English proficiency	.25	.19	.04	4.47	<.001
Education	.26	.28	.06	4.54	<.001
Age	–.18	–.95	.30	–3.18	.001
Female	–.15	–.05	.02	–2.73	.01

Discussion

Principal Findings

In this study, we examined the association of PU, PEOU, ICT use, and loneliness among low income, predominately Asian American, older adults. This research built upon 2 previous studies that simplified the TAM to examine demographic and health factors that impact technology acceptance among Asian Americans aged 62 years and older living in affordable senior housing communities [17,18]. In these previous analyses, the conceptualization of PU, PEOU, and ICT use were based on the existing literature and theory around how these constructs apply to older adults. The constructs were developed by summing and normalizing 12 survey items. In this study, the assumptions underlying the conceptualization of PU, PEOU, and ICT use and the TAM framework were rigorously assessed and confirmed using a combination of statistical techniques. Initially, EFA was conducted to determine the appropriate number of factors and explore the underlying structure of the constructs. Horn parallel analysis was subsequently used to validate the EFA results. Following this, SEM (beginning with CFA) was performed to validate and confirm the factor structure identified in EFA and examine the relationships among PU, PEOU, ICT use, and loneliness.

The original TAM focused on technology acceptance in the workplace. PU was measured with items such as, "Using [chart master] in my job would enable me to accomplish tasks more quickly," and, "Using [chart master] would improve my job performance" [21]. Our research findings support a modified construct of PU that accounts for older adults' perceptions of ICTs as being useful for social connection and for learning new information and skills. We found that the item, "Technology helps me learn new information and skills," had a slightly higher factor loading than the item, "Technology helps me be connected with family and friends." Mixed methods longitudinal data from the Lighthouse Project for Older Adults supported this finding; participants reported that they most frequently used ICTs to access YouTube (eg, to view videos related to nutrition, exercise, and cultural content), followed by accessing entertainment and checking the weather [32]. The Pew Research Center reported that, in 2021 among adults aged 65 years and older, the use of YouTube experienced the most growth as compared with any other app [33]. Notably, studies suggest that app- and web-based activities among older adults vary significantly by age group and gender [23,34-36].

When it comes to PU, it is essential to note that the COVID-19 pandemic spurred an exponential increase in the use of ICTs for social connection. American Association of Retired Persons (AARP) 2021 Tech Trends reported a notable surge in the use of various communication technologies among individuals aged 50 years and above to stay connected with others. A significant portion of this demographic reported an uptick in their usage of video chats (45%, 1022/2271), texting (37%, 840/2271), emailing (26%, 590/2271), and phone calls (29%; 659/2271) compared with prepandemic levels. In 2019, approximately half had never used video chat, whereas by 2020, this figure rose to 70%, with one out of 3 engaging in video chats on a weekly basis [37].

Our findings supported the operationalization of PEOU combining 2 items from the Senior TAM – Tech Anxiety Subscale [24] and 3 items from the ATCQ – Comfort Subscale [25]. Previous studies note an array of emotions that influence technology acceptance and use, including enjoyment, effort expectancy [38], control, efficacy [24,39], confidence [40,41], comfort [25,42,43], and anxiety [24,40]. Due to the variety of existing measures and constructs highlighted in the literature (some developed in the 1980s and with highly educated, White, middle-aged adults), it can be challenging for researchers to select a concise set of items that are specific to older adults and modern technology. We believe our findings can inform future studies with Asian American older adults, who require survey modification [44]. During exploratory analysis, one item, "I feel comfortable with technology," was dropped due to low factor loading. We believe this may have been due to the item being reverse scored, which may have been confusing due to participants' limited literacy levels. Our findings align with previous research which suggests that assessment scales containing reverse-scored items impose higher cognitive processing demands, potentially resulting in measurement challenges for older adult participants [45], particularly since our participants had self-reported limited English proficiency and low educational attainment (Table 1).

Our mediating factor, ICT use, was unique in that it combined measures assessing the frequency of use of smartphones, tablets, and computers, as well as years of experience. We believe this measure is valuable to better understand ICT use since these devices are often used interchangeably among older adults for multitasking [46]. Furthermore, we are not the first to identify years of experience as an important factor in understanding technology acceptance [38].

Using CFA, we found significant, positive relationships between PU, PEOU, and ICT use. This finding is aligned with the Senior Technology Exploration, Learning, and Acceptance model [47] and the Senior Technology Acceptance and Adoption Model (STAM) [48]. The concept of reinforcement of use is consistent with a model of technology acceptance or rejection from an ease-of-learning perspective [49]. When ICTs are perceived as more useful, their usage tends to increase, forming a reinforcing cycle. Increased usage and familiarity make it easier, encouraging individuals to master new skills and diversify their usage, such as progressing from social chats to watching YouTube videos, further reinforcing their mastery and use.

SEM confirmed our modified TAM, supporting our three hypotheses. PEOU significantly, and positively predicted PU and ICT use. Further, ICT use was significantly negatively associated with loneliness. These relationships remained significant, even when adjusting for gender, age, education, English proficiency, Asian ethnicity, and subjective cognitive decline. Interestingly, English proficiency was related to PEOU and ICT use, suggesting a cultural or linguistic bias in technology and application development favoring English-speaking users. Asian ethnicity was associated with loneliness, affirming observations of the vital role technology

can play for immigrants in maintaining social and cultural connections [11]. Subjective cognitive decline was also associated with loneliness, perhaps a reflection of lower social and technology engagement [50].

This study has implications for interventions, particularly among populations with lower literacy. Both PEOU and PU are modifiable factors and could be enhanced by offering digital literacy training and support to new learners [43,51-53] or tailoring user interfaces [54,55]. The relevance of technology for the individual can be increased by demonstrating various culturally relevant use cases for devices, such as connecting with distant relatives, accessing health information, or enjoying entertainment in one's native language. In the Lighthouse Project for Older Adults, the low-income housing providers made a commitment to enhance technology use and offered the technology, training, and support to encourage adoption. This not only provided tangible support in the form of equipment and training but also created a community of learners where individuals could benefit from the experience of peers [32]. This is an example of how service providers can play a role in potentially improving quality of life and reducing loneliness by encouraging the adoption of technology.

Study Limitations

This study was partially limited by the use of a convenience sample of participants who were interested in learning more about technology, not a random sample. This potentially biases the sample toward those with more positive attitudes toward technology. Even with this bias, we observed variability in perceptions about technology but did not capture the full array of attitudes likely present in the population. In addition, participants all lived in age-restricted affordable housing communities. Therefore, our findings are not representative of older adults living in multigenerational households. We removed participants whose data were missing the dependent variable, loneliness. We retained all other participants, yet 4%-7% of responses were missing for items related to PEOU and PU.

Methodological considerations included minimizing participant burden to maximize potential engagement with technology and the most complete dataset. This required modifications to measures, to accomplish parsimony and ease of administration, potentially compromising psychometric properties. However, our CFA indicated the adequacy of the measures for crucial constructs, despite having fewer than 3 observations for each latent variable in the case of PU.

EFA and SEM were limited by the inclusion of binary and ordinal observed variables (eg, gender, Asian ethnicity, and subjective cognitive decline). While this represents a methodological weakness, these variables were included because of their known importance to the constructs of interest. Furthermore, we used the same dataset for all multivariate analyses because of the early phase of understanding relationships among our variables of interest, recognizing that typically the confirmatory analysis should be conducted on an independent sample.

Finally, many other factors potentially impact technology acceptance among older adults, including social and health factors, access to Wi-Fi and devices, and digital literacy training.

Conclusion

Despite the limitations, this study affirms the usefulness of the TAM in understanding the dynamics of technology adoption among a low-income Asian American population. At baseline, there is considerable interest in technology, affirming its relevance in the lives of older adults. The role of English proficiency in ICT use warrants further exploration to identify ways to increase equity and access for those who have another primary language. The study further highlights the potential role that technology could play in alleviating loneliness through greater engagement with family and friends and the ability to maintain cultural ties. Future studies could explore the most effective ways to overcome resistance to technology, the most meaningful ways to support novice users to adopt a new device, and ways to increase the diversity of use once a user has become comfortable with basic functions. While technology cannot replace the human touch, it has the promise to improve engagement and social connection among isolated, low-income older adults.

Acknowledgments

This work was funded by the Healthy Aging in a Digital World Initiative at the University of California, Davis and the Center for Information Technology Research in the Interest of Society (CITRIS) and the Banatao Institute.

Other acknowledgments: The University of California, Davis, School of Medicine Office of Research, Evaluation Unit, including Rebeca Giacinto, Pauline DeLange Martinez, Stuart Henderson, and Julie Rainwater, designed the Lighthouse surveys and provided oversight for data collection and entry. We thank all Lighthouse participants for their time and effort in completing the evaluation surveys. We also thank the staff at each site for their support with participant recruitment, data collection, and entry.

Lighthouse project leadership included the Front Porch Center for Innovation and Wellbeing (Davis Park, Kari Olson, Jennifer Lee), Eskaton (Therese ten Brinke, Sheri Peifer), and CITRIS (David Lindeman, Gale Berkowitz).

Data Availability

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

Authors' Contributions

PDM led conceptual design, data analysis, and authorship. HMY made substantial conceptual and design contributions, providing invaluable input along all stages of manuscript preparation. MP and DJT provided oversight for the operationalization of key variables, data analysis, reporting results, and interpretation of the data. LG helped to critically revise the manuscript and agreed to be accountable for the accuracy of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Lighthouse survey.
[DOCX File , 53 KB - aging_v8i1e63856_app1.docx]

Multimedia Appendix 2
Correlation matrix.
[XLSX File (Microsoft Excel File), 20 KB - aging_v8i1e63856_app2.xlsx]

References

1. UCLA Center for Health Policy Research. AskCHIS. Loneliness Scale. Los Angeles. URL: https://healthpolicy.ucla.edu/our-work/askchis [accessed 2024-11-20]
2. New Surgeon General Advisory Raises Alarm about the Devastating Impact of the Epidemic of Loneliness and Isolation in the United States. 2023. URL: https://www.hhs.gov/about/news/2023/05/03/new-surgeon-general-advisory-raises-alarm-about-devastating-impact-epidemic-loneliness-isolation-united-states.html [accessed 2024-01-28]
3. Shimkhada R, Tse HW, Ponce NA. Life satisfaction and social and emotional support among asian American older adults. J Am Board Fam Med 2022;35(1):203-205 [FREE Full text] [doi: 10.3122/jabfm.2022.01.210232] [Medline: 35039430]
4. Malani P, Singer D, Kirch M, Solway E, Roberts S, Smith E. Trends in loneliness among older adults from 2018-2023. University of Michigan National Poll on Healthy Aging; 2023 updated March 2019;3(1):600-601 [FREE Full text] [doi: 10.1093/geroni/igz038.2233]
5. Pinquart M, Sörensen S. Risk Factors for Loneliness In Adulthood and Old Age--A Meta-Analysis. Hauppauge, New York: Nova Science Publishers; 2003.
6. Cohen-Mansfield J, Hazan H, Lerman Y, Shalom V. Correlates and predictors of loneliness in older-adults: a review of quantitative results informed by qualitative insights. Int. Psychogeriatr 2015;28(4):557-576. [doi: 10.1017/s1041610215001532]
7. Delaruelle K. Migration-related inequalities in loneliness across age groups: a cross-national comparative study in Europe. Eur J Ageing 2023;20(1):35 [FREE Full text] [doi: 10.1007/s10433-023-00782-x] [Medline: 37612470]
8. Wang Y, Woo B, Jiang N. Correlates of loneliness among older asian immigrants: a systematic review. Innovation in Aging 2019;3(1):629. [doi: 10.1093/geroni/igz038.2345]
9. Petersen B, Khalili-Mahani N, Murphy C, Sawchuk K, Phillips N, Li KZH, et al. The association between information and communication technologies, loneliness and social connectedness: a scoping review. Front Psychol 2023;14:1063146 [FREE Full text] [doi: 10.3389/fpsyg.2023.1063146] [Medline: 37034933]
10. Ibarra F, Baez M, Cernuzzi L, Casati F. A systematic review on technology-supported interventions to improve old-age social wellbeing: loneliness, social isolation, and connectedness. J Healthc Eng 2020;2020:2036842 [FREE Full text] [doi: 10.1155/2020/2036842] [Medline: 32765823]
11. Khvorostianov N, Elias N, Nimrod G. 'Without it i am nothing': the internet in the lives of older immigrants. New Media & Society 2011;14(4):583-599. [doi: 10.1177/1461444811421599]
12. Melenhorst AS, Rogers WA, Caylor EC. The use of communication technologies by older adults: exploring the benefits from the user's perspective. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2001;45(3):221-225. [doi: 10.1177/154193120104500305]
13. Blacher K. National asian pacific center on aging. Asian Americans and Pacific Islanders in the United States Aged 65 Years and Older: Population, Nativity, and Language. 2013. URL: https://www.napca.org/wp-content/uploads/2017/10/65-population-report-FINAL.pdf [accessed 2021-12-30]
14. Yoon H, Jang Y, Kim S, Speasmaker A, Nam I. Trends in internet use among older adults in the United States, 2011-2016. J Appl Gerontol 2021;40(5):466-470. [doi: 10.1177/0733464820908427] [Medline: 32131670]
15. Yoon H, Jang Y, Vaughan PW, Garcia M. Older adults' internet use for health information: digital divide by race/ethnicity and socioeconomic status. J Appl Gerontol 2020;39(1):105-110. [doi: 10.1177/0733464818770772] [Medline: 29661052]
16. Kim J, Lee HY. Older minority immigrants' digital technology use. Does English Proficiency Matter? The Gerontologist 2016;56(3):139. [doi: 10.1093/geront/gnw162.557]

17. DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Heather MY. The role of health in the technology acceptance model among low-income, asian American older adults: a cross sectional survey analysis. JMIR Formative Research 2024:TBD (forthcoming) [FREE Full text] [doi: 10.2196/57009]
18. Delange Martinez P, Tancredi DJ, Pavel M, Garcia L, Young HM. JMIR. Technology Acceptance Among Low-In come, Asian American Older Adults: Cross-Sectional Survey Analysis 2024 Nov 22;26:e52498 [FREE Full text] [doi: 10.2196/52498] [Medline: 39576987]
19. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564 [FREE Full text] [doi: 10.2196/44564] [Medline: 37191976]
20. Charness N, Boot WR. Chapter 20 - technology, gaming, and social networking. In: Schaie KW, Willis SL, editors. Handbook of the Psychology of Aging (Eighth Edition). Saon Dieg: Academic Press; 2016.
21. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319. [doi: 10.2307/249008]
22. CITRIS and the Banatao Institute. Brings technology-enabled health and well-being to low-income california seniors during COVID crisis. Lighthouse for Older Adults, a CITRIS and University of California Initiative. 2020. URL: https://citris-uc.org/lighthouse-for-older-adults/ [accessed 2020-09-16]
23. Sims T, Reed AE, Carr DC. Information and communication technology use is related to higher well-being among the oldest-old. J Gerontol B Psychol Sci Soc Sci 2017;72(5):761-770. [doi: 10.1093/geronb/gbw130] [Medline: 27702839]
24. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
25. Jay GM, Willis SL. Influence of direct computer experience on older adults' attitudes toward computers. J Gerontol 1992;47(4):250-257. [doi: 10.1093/geronj/47.4.p250] [Medline: 1624702]
26. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672 [FREE Full text] [doi: 10.1177/0164027504268574] [Medline: 18504506]
27. Tabachnick BG, Fidell LS. In: Boston MA, editor. Using Multivariate Statistics. USA: Pearson Education, Inc; 2018.
28. Kaplan D. Structural equation modeling. In: Smelser NJ, Baltes PB, editors. International Encyclopedia of the Social & Behavioral Sciences. Oxford: Pergamon; 2001:15215-15222.
29. Hair JFJ, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate Data Analysis. New Jersey: Prentice-Hall; 2009.
30. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 1981;18(1):39. [doi: 10.2307/3151312]
31. Hair JFJ, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate Data Analysis. New Jersey: Prentice-Hall; 2018.
32. Giacinto RE, Martinez P, Neves S, Sullivan M, Kipke R, Henderson S. University of california, davis, school of medicine office of research evaluation unit,. Lighthouse for Older Adults Evaluation Report. URL: https://health.citris-uc.org/wp-content/uploads/2023/08/Lighthouse-UCD-SOMOR-Evaluation-Summary-Report-FINAL-PDF.pdf [accessed 2024-11-20]
33. Faverio M. Pew Research Center. Share of those 65 and older who are tech users has grown in the past decade. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2022-01-13]
34. Anderson GO. Getting Connected Older Americans Embrace Technology to Enhance Their Lives. AARP Research. 2018. URL: https://www.aarp.org/research/topics/technology/info-2018/technology-use-attitudes.html [accessed 2018-02-23]
35. Kim J, Lee HY, Christensen MC, Merighi JR. Technology access and use, and their associations with social engagement among older adults: do women and men differ? J Gerontol B Psychol Sci Soc Sci 2017;72(5):836-845. [doi: 10.1093/geronb/gbw123] [Medline: 28073816]
36. Yu RP, Ellison NB, McCammon RJ, Langa KM. Mapping the two levels of digital divide: internet access and social network site adoption among older adults in the USA. Information, Communication & Society 2015;19(10):1445-1464. [doi: 10.1080/1369118x.2015.1109695]
37. Kakulla B. Personal Tech and the Pandemic.: Older Adults Are Upgrading for a Better Online Experience. AARP Research URL: https://www.aarp.org/research/topics/technology/info-2021/2021-technology-trends-older-americans.html-CMP=RDRCT-PRI-TECH-040721/?cmp=RDRCT-907b618d-20210416 [accessed 2024-02-05]
38. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
39. Cassidy S, Eachus P. Developing the computer user self-efficacy (Cuse) scale: investigating the relationship between computer self-efficacy, gender and experience with computers. Journal of Educational Computing Research 2002;26(2):133-153. [doi: 10.2190/jgjr-0kvl-hrf7-gcnv]
40. Loyd BH, Gressard C. Reliability and factorial validity of computer attitude scales. Educational and Psychological Measurement 1984;44(2):501-505. [doi: 10.1177/0013164484442033]
41. Nickell GS, Pinto JN. The computer attitude scale. Computers in Human Behavior 1986:301-306 [FREE Full text] [doi: 10.1016/0747-5632(86)90010-5]

42. Laganà L, Oliver T, Ainsworth A, Edwards M. Enhancing computer self-efficacy and attitudes in multi-ethnic older adults: a randomised controlled study. Ageing Soc 2011;31(6):911-933 [FREE Full text] [doi: 10.1017/S0144686X10001340] [Medline: 25512679]
43. Arthanat S. Promoting information communication technology adoption and acceptance for aging-in-place: a randomized controlled trial. J Appl Gerontol 2021;40(5):471-480 [FREE Full text] [doi: 10.1177/0733464819891045] [Medline: 31782347]
44. Stewart AL, Thrasher AD, Goldberg J, Shea JA. A framework for understanding modifications to measures for diverse populations. J Aging Health 2012;24(6):992-1017 [FREE Full text] [doi: 10.1177/0898264312440321] [Medline: 22495768]
45. Carlson M, Wilcox R, Chou CP, Chang M, Yang F, Blanchard J, et al. Psychometric properties of reverse-scored items on the CES-D in a sample of ethnically diverse older adults. Psychol Assess 2011;23(2):558-562 [FREE Full text] [doi: 10.1037/a0022484] [Medline: 21319906]
46. Kononova A, Joshi P, Cotten S. Contrary to myth, older adults multitask with media and technologies, but studying their multitasking behaviors can be challenging. Innov Aging 2019;3(4):igz029 [FREE Full text] [doi: 10.1093/geroni/igz029] [Medline: 31637313]
47. Tsai HYS, Rikard RV, Cotten SR, Shillair R. Senior technology exploration, learning, and acceptance (STELA) model: from exploration to use – a longitudinal randomized controlled trial. Educational Gerontology 2019;45(12):728-743. [doi: 10.1080/03601277.2019.1690802]
48. Renaud K, Biljon JV. Predicting technology acceptance and adoption by the elderly: a qualitative study. 2008 Presented at: SAICSIT '08: Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology; 2008 October 8; New York. [doi: 10.1145/1456659.1456684]
49. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 2013;29(4):1715-1724 [FREE Full text] [doi: 10.1016/j.chb.2013.02.006]
50. Li Y, Godai K, Kido M, Komori S, Shima R, Kamide K, et al. Cognitive decline and poor social relationship in older adults during COVID-19 pandemic: can information and communications technology (ICT) use helps? BMC Geriatr 2022;22(1):375 [FREE Full text] [doi: 10.1186/s12877-022-03061-z] [Medline: 35484494]
51. Lagana L, Garcia JJ. The mental health impact of computer and internet training on a multi-ethnic sample of community-dwelling older adults: results of a pilot randomised controlled trial. Int. J Biomed Sci 2013;9(3):135-147 [FREE Full text] [doi: 10.59566/ijbs.2013.9135]
52. González A, Ramírez MP, Viadel V. ICT learning by older adults and their attitudes toward computer use. Curr Gerontol Geriatr Res 2015;2015:849308 [FREE Full text] [doi: 10.1155/2015/849308] [Medline: 26346158]
53. Senteio CR. Investigating the enduring impact of a community-based health education program to promote African American elders' use of technology designed to support chronic disease self-management. Geriatrics (Basel) 2018;3(4):70 [FREE Full text] [doi: 10.3390/geriatrics3040070] [Medline: 31011105]
54. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the PRISM randomized controlled trial. Gerontologist 2018;58(3):467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
55. Lor M, Bowers BJ. Feasibility of audio-computer-assisted self-interviewing with color-coding and helper assistance (ACASI-H) for hmong older adults. Res Nurs Health 2017;40(4):360-371 [FREE Full text] [doi: 10.1002/nur.21796] [Medline: 28431187]

Abbreviations

ATCQ: Attitudes Towards Computers Questionnaire
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: Comparative Fit Index
CHIS: California Health Interview Survey
CMIN: chi-square minimum
EFA: exploratory factor analysis
ICT: Information communication technology
NHATS: National Health and Aging Trends Study
PEOU: perceived ease of use
PU: perceived usefulness
RMSEA: root-mean-square error of approximation
SEM: structural equation modeling
STAM: Senior Technology Acceptance and Adoption Model
TAM: Technology Acceptance Model

TLI: Tucker-Lewis Index

Edited by PH Chau; submitted 01.07.24; peer-reviewed by R Shimkhada, Z Luo; comments to author 12.09.24; revised version received 22.10.24; accepted 06.11.24; published 08.01.25.

Please cite as:
DeLange Martinez P, Tancredi D, Pavel M, Garcia L, Young HM
Adapting the Technology Acceptance Model to Examine the Use of Information Communication Technologies and Loneliness Among Low-Income, Older Asian Americans: Cross-Sectional Survey Analysis
JMIR Aging 2025;8:e63856
URL: https://aging.jmir.org/2025/1/e63856
doi:10.2196/63856
PMID:

Original Paper

Identifying Food Preferences and Malnutrition in Older Adults in Care Homes: Co-Design Study of a Digital Nutrition Assessment Tool

Jenni Connelly[1*], BA, MSc, PhD; Kevin Swingler[2], BSc, MSc, PhD; Nidia Rodriguez-Sanchez[1], BSc, PhD; Anna C Whittaker[1*], BSc, MSc, PhD

[1]Faculty of Health Sciences and Sport, University of Stirling, Stirling, United Kingdom

[2]Faculty of Natural Sciences, University of Stirling, Stirling, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Jenni Connelly, BA, MSc, PhD
Faculty of Health Sciences and Sport
University of Stirling
3a74a Cottrell Building
Stirling, FK9 4LA
United Kingdom
Phone: 44 1786 466399
Email: jenni.connelly1@stir.ac.uk

Abstract

Background: Malnutrition is a challenge among older adults and can result in serious health consequences. However, the dietary intake monitoring needed to identify malnutrition for early intervention is affected by issues such as difficulty remembering or needing a dietitian to interpret the results.

Objective: This study aims to co-design a tool using automated food classification to monitor dietary intake and food preferences, as well as food-related symptoms and mood and hunger ratings, for use in care homes.

Methods: Participants were 2 separate advisory groups and 2 separate sets of prototype testers. The testers for the first prototype were 10 community-dwelling older adults based in the Stirlingshire area in Scotland who noted their feedback on the tool over 2 weeks in a food diary. The second set of testers consisted of 14 individuals (staff: n=8, 57%; and residents: n=6, 43%) based in 4 care homes in Scotland who provided feedback via interview after testing the tool for a minimum of 3 days. In addition, 130 care home staff across the United Kingdom completed the web-based survey on the tool's needs and potential routes to pay for it; 2 care home managers took part in follow-up interviews. Data were collected through food diaries, a web-based survey, audio recordings and transcriptions of focus groups and interviews, and research notes. Systematic text condensation was used to describe themes across the different types of data.

Results: Key features identified included ratings of hunger, mood, and gastrointestinal symptoms that could be associated with eating specific foods, as well as a traffic light system to indicate risk. Issues included staff time, Wi-Fi connectivity, and the accurate recognition of pureed food and fortified meals. Different models for potential use and commercialization were identified, including peer support among residents to assist those considered less able, staff-only use of the tool, care home–personalized database menus for easy meal photo selection, and targeted monitoring of residents considered to be at the highest risk using the traffic light system.

Conclusions: The tool was deemed useful for monitoring dietary habits and associated symptoms, but necessary design improvements were identified. These should be incorporated before formal evaluation of the tool as an intervention in this setting. Co-design was vital to help make the tool fit for the intended setting and users.

(JMIR Aging 2025;8:e64661) doi:10.2196/64661

KEYWORDS

ageing; digital technology; dietary measurement; care homes; co-design; dietary intake; food diary

Introduction

Background

Diet is an important component of a healthy lifestyle, and it remains one of the most common challenges among older adults in the United Kingdom [1]. Malnutrition is also a prevalent issue in this population. The World Health Organization defines malnutrition as "deficiencies or excesses in nutrient intake, imbalance of essential nutrients or impaired nutrient utilization" [2]. Approximately 10% of those aged 65 years are malnourished or are at risk of malnutrition in the United Kingdom [3], a figure that rose to 25% of those aged 60 years during the COVID-19 pandemic [4-6]. The prevalence of severe malnutrition is reported to be as high as 45% in those living in residential care homes [7] but may be worse because approximately 70% of malnutrition in the United Kingdom is unrecognized [8,9]. Malnutrition is strongly associated with frailty, as well as functional and health decline [10-12]; thus, it is important to monitor dietary intake and identify malnutrition risk before symptoms are already evident. However, tracking and understanding the impact of dietary intake is not easy. It is possible for people to become malnourished if they do not eat enough nutritionally rich foods for 2 to 5 days [6,13].

Malnutrition in older adults is commonly identified via the Malnutrition Universal Screening Tool (MUST) [14], and when individuals are classed as medium to high risk, it is recommended to document food and drink intake for 3 days. However, there are difficulties with self-reported food diaries, which can be inaccurate due to recording error, misremembering, and socially desirable responses. Additional barriers include limited staff time in care homes and the need for analysis and interpretation by a dietitian [15].

Computer-based tailored interventions may offer a solution because there is evidence that they have positively impacted dietary monitoring in the general population [16]. Automated food classification is a rapidly advancing field of artificial intelligence research that uses a computer source to identify different foods from an image [17,18]. Creating new nutrition databases or using or adapting existing databases means that it is possible to infer detailed nutritional data from photos of food [19]. Convolutional neural network algorithms are used to identify food automatically from photos, and these are linked to a nutritional database to calculate the calories in the identified foods. Machine learning techniques can then be applied to correlate dietary intake with important individual differences such as food preferences, physical symptoms, and malnutrition risk [20]. Automated food classification can consider the portion size and regularity of food intake, which is crucial to monitor in older people because this can significantly influence malnutrition risk. Linking food intake to preferences and symptoms through technology and artificial intelligence means this information could be used to identify potential risks quickly and accurately. This would provide care staff with information to help monitor older people's diet and decrease malnutrition risk. Studies that have evaluated photo-based dietary assessment tools have found measurement errors comparable to traditional methods [21], and those that have found it a valid method have focused on a younger population, reporting positive feedback [22,23].

Objectives

Consequently, following the 6 steps for quality intervention development (6SQuID) [24], we sought to co-design a prototype tool with key stakeholders, which we refer to as a "digital nutritional assessment tool." We aimed to work with advisory groups (AGs) and participants to develop the key features of a prototype tool for monitoring dietary intake, map this onto food preferences and symptoms, and test the prototype with stakeholders. It is important to note that this study reports on developmental research to identify the tool's suitability, features needed, and usability, rather than its efficacy at capturing and estimating nutrient intake, which is part of the next stage of development and testing. Its key objectives were to complete the first two 6SQuID steps: (1) defining and understanding the problem and its causes and (2) identifying which causal or contextual factors are modifiable—which have the greatest scope for change and who would benefit the most—using a co-design and coproduction approach. This inclusive approach was taken due to the wish for the resulting tool to have real-life application and translation outside of the academic setting while acknowledging the importance of specifying up front the motivations of coproduction, what outcomes are required for whom, and how these might be achieved [25]. This qualitative approach to data collection is necessarily reflective and conscious and has attempted to follow recently produced resources as a guide to quality coproduction methods in health research [26,27]. In this way, it positions end users as essential to the research process; thus, it is collaborative and equitable [27]. In this study, stakeholders were identified as older adults, the tool software designers, dietitians, and ultimately care home staff and residents themselves.

Methods

Recruitment

Participants consisted of 2 AGs and 2 separate sets of prototype testers. Both AGs comprised different older adults and dietitians. Care home staff as well as app and tool designers were consistent across the 2 groups. The testers for the first prototype were community-dwelling older adults and care home employees recruited across Stirlingshire in Scotland. For care home testing, the aim was to recruit 3 care homes, 2 staff members, and 4 to 5 residents in each care home. Care home staff and managers were recruited via social media and Scottish Care newsletters to complete a survey and semistructured interviews, respectively.

Study Design

The study used a co-design and coproduction approach [27] to determine key components of the tool to be developed and used through multiple iterations of usability feedback from older adults, care staff, and dietitians. The tool was developed using the information provided by the AGs in a collaboration between the research team and the app developers, Game Doctor. The development process involved AGs with key stakeholders in the design process (phase 1); community testing and feedback (phase 2); and care setting feasibility and usability testing, as

well as consultation and data gathering with a wider group of care home staff and managers (phase 3). This was a continual process, with feedback resulting in a new build of the tool prototype, to 3 builds. A summary of the timeline for the study design and the different groups involved is shown below in Table 1.

Table 1. Timeline and summary of the study design.

	Phase 1			Phase 2							Phase 3				
	Month 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AG[a] 1		✓		✓											
Testing				✓				✓				✓			
Survey									✓	✓				✓	
Interviews										✓			✓		
AG 2											✓				✓

[a]AG: advisory group.

Procedure

Phase 1: AG Recruitment and Co-Design of Tool

According to the 6SQuID framework [24], the first step of intervention development is "defining and understanding the problem." This was achieved through the creation of the first AG (AG 1) with key stakeholders to develop ideas and identify features that would be vital for a food recognition tool in the care home setting. The study and AG role were advertised through email to the Stirling 1000 Elders email list used by ACW to recruit 2 older adults. A snowball technique was then used to recruit key stakeholders working within nutrition and care settings. Through this process, AG 1 with 6 key stakeholders (older adults: n=2, 33%; care home staff: n=1, 17%; app developers: n=2, 33%; and dietitian: n=1, 17%) was formed. The first session was run face-to-face and was used to define and understand the problem (ie, the need for a food recognition model for dietary monitoring and how artificial intelligence could be used as a potential solution). In addition, AG 1 members were asked to brainstorm key features that the tool should include, such as the identification of symptoms related to eating a particular food and potential rating scales to assess mental and physical well-being after eating. AG 1 also identified potential issues and challenges that would be faced in using the tool with older adults in a health care setting generally.

Feedback from this first AG 1 meeting provided the app developers, Game Doctor, with information on which to base the initial build of the tool and the 2 subsequent builds based on later AG meetings. The LogMeal (AIGecko Technologies SL, Artificial Intelligence and Deep Learning Food Division) photo database used by the tool to identify the foods photographed was selected in collaboration between the research team and app developers based on a review of potential existing services, considering the range of foods recognized and financial costs. The first prototype was then developed by Game Doctor, linked to the photo database, and demonstrated in a second meeting of AG 1 for prototype testing and feedback before bug fixing and independent feasibility testing.

Phase 2: Feasibility Testing of the Tool and Food Diaries

This phase was based on 6SQuID step 2: "identify which causal or contextual factors are modifiable—which have the greatest scope for change and who would benefit the most." We recruited 10 community-dwelling older adults via email to Stirling 1000 Elders. Eligibility included owning an Android mobile phone or tablet capable of downloading the tool. The older adults were provided with a pseudonymized log-in ID for the tool and asked to use it to record each meal over 2 weeks by photographing their plate before and after eating. They were also asked to select the foods present that were identified by the tool or to correct any misidentified foods by entering what the foods circled in the tool were. They could then also rate on a 5-point emoji-based Likert scale their mood, level of hunger after eating, and the presence of up to 7 further symptoms. The symptoms included in the tool were based on feedback from AG 1 and included tiredness, diarrhea, cramping, swallowing difficulties, constipation or discomfort, and bloating. A written diary to complete alongside using the tool was provided to participants to indicate the contents of each meal and note down any problems with using the tool and provide general feedback about it. A third and final meeting of AG 1 was planned to discuss the findings from prototype testing and give the members the opportunity for further testing to provide feedback to the app developer so that a third version of the prototype could be developed. Due to difficulties in arranging a final meeting, AG 1 members were asked to use the tool for a week and then complete a short web-based survey. The survey asked about missing tool features and symptoms, strategies to support the tool's integration in care settings, and other populations that would benefit from using the tool. This was then used as the basis for discussions among members of the second AG (AG 2).

Phase 3

Overview

As the tool was intended for eventual use in care homes, phases 1 and 2 were repeated with AG 2, a newly recruited group of 4 key stakeholders (older adults: n=2, 50%; care home staff: n=1, 25%; and dietitian [also an older adult]: n=1, 25%). In addition,

a new cohort of tester participants, including staff and residents, was recruited through 4 care homes.

At the first AG 2 meeting, the members discussed version 3 of the prototype and the challenges of its application in care homes.

Care Home Testing and Focus Groups

For testing, 8 care homes were contacted through existing networks with information on the study and what was expected of staff and residents. The lead researcher (JC) prepared an introduction to the tool and led a separate discussion at each care home on key features that care home staff and residents wanted to see in this type of tool. Each care home then had access to version 3 of the prototype for 2 weeks, and staff and residents were asked to record food intake and food-related symptoms. The care homes were given an Android tablet preloaded with the tool that could be used to take photos of the food. During this period, residents (n=6) were asked to take photos of all foods they consumed, while staff members (n=8) recorded food intake for other residents, across a minimum of 3 days, ensuring that each type of meal (breakfast, lunch, dinner, and snacks) was recorded at least once. The photos were then added to a timeline, and participants had the option to indicate how hungry they felt after the meal using an emoji-based Likert scale and click on any symptoms they were experiencing from a list of common symptoms, which were then added to the timeline. Participants also entered any verbal preferences and indicated whether they needed support with eating.

After prototype testing, JC conducted semistructured interviews with 5 participants (care home staff: n=4, 80%; and resident: n=1, 20%) to obtain their feedback on the usability and acceptability of the tool. Questions covered topics such as what they liked about the tool, difficulties encountered with its use, missing features (eg, physical or mental symptoms that should be included), additional features that should be included, and advice on the usability of the tool in practice (eg, who would take the photos and whether there was enough time to use the tool during a typical care home day). A second and final AG 2 meeting was convened to talk through the feedback from care homes and suggest future directions.

Care Home Staff Perspectives Survey and Manager Interviews

Concurrently with the care home testing, a web-based survey was devised for care home staff that included 16 questions under three headings: (1) “About you” (including job title, organization, and time spent in the care sector), (2) “Nutrition” (including questions about monitoring diet and dietary outcomes in their care home organization), and (3) “Routes to integrate tool.” Of the 16 questions, 6 (38%) were open ended. These were included (under the appropriate headings listed previously) to explore any care home–specific strategies or features that would improve the integration of the tool. The survey was designed to take no longer than 20 minutes to complete. This included a 7-minute video on the web showcasing the tool and its current features, which was required to be watched before completing the survey.

A total of 130 participants (care home staff) were recruited across Scotland via social media (ie, Twitter [subsequently rebranded as X] and LinkedIn), emails to existing partner care homes, and through Scottish Care. The inclusion criteria included any worker from within a UK care home. The survey was available through the advertising link, and, when clicked, the individual was taken to the participant information sheet.

After the survey, semistructured interviews were conducted with care home managers to explore the survey feedback and assess how the tool would fit within their care homes. Care home managers could sign up for these interviews via the last question in the survey.

Data Analysis

This project used systematic text condensation [28] across the different types of qualitative data collected. This is a descriptive and explorative method for gathering themes across different types of qualitative data, including interviews, observations, and written text analysis. As such, it represents a pragmatic approach that can turn total impressions of the overall chaos of the data into themes with units of meaning that can then be condensed and synthesized into descriptions and concepts.

Ethics Approval

This study was conducted in accordance with the Declaration of Helsinki and approved by the General University Ethics Committee of the University of Stirling (8857; September 7, 2023) for studies involving human participants. As a token of thanks, the participants testing the app received £50 (US $63.32) at the end, once the device had been returned. The feedback provided on the tool was anonymized in terms of users but not per care home. This was carried out in this manner to ensure feedback could be related to the size and location of the care home. Those who took part in the advisory group received £50 (US $63.32) per meeting they attended. For both the survey and the follow-up interviews, participants implied consent by clicking the provided link to take part in the survey.

Results

Participants

Participants consisted of 2 separate AGs (AG 1 and AG 2) and 2 separate sets of prototype testers. The testers for the first prototype were 10 community-dwelling older adults and care home employees based in the Stirlingshire area. Four care homes responded about participating in the study. The care home testers consisted of 14 individuals (staff: n=8, 57%; and residents: n=6, 43%) in 4 care homes across Crieff, Stirling, and Fife, differing in size and type of organization (nursing care, older adults only, mixed needs, and dementia specialty or general care). Altogether, 130 individuals completed the web-based survey, with 2 managers volunteering to take part in follow-up interviews.

Phase 1: AG Recruitment and Co-Design of Tool

AG 1 members met with the research team for 2 hours on 2 occasions over the course of 1 year to design, test, and improve the first protype. They also tested the final prototype and gave feedback before testing in community-dwelling older adults.

Game Doctor staff took part in all AGs and produced an initial prototype in March 2022. Game Doctor. They suggested that the 3-sprint model would work well for the development of up to 3 builds of the initial prototype within the project timeline and budget. At the first meeting of AG 1, key features for the prototype were discussed, and priority was assessed. The need for a food recognition model and key features it should have, such as simple and quick food recognition, were noted. How it would work in the care home environment, issues of data security and consent, symptoms that needed to be included, potential for a mood scale, and overall usability of the tool were also considered. Priorities were summarized into a set of features to be incorporated (Table 2).

Table 2. Key features identified through the advisory group, along with the priority level determined in collaboration with Game Doctor.

Feature	ID	Description	Priority
Food tracking via image recognition model	F1V1	Food tracking using off the shelf image recognition model	High
Patient profiles	F2V1	A way to create profiles for each patient or resident using the application	High
Correct food button input	F3V1	Input button to confirm that the model has recognized food accurately	High
Before or after plate recognition	F4V1	Ability of model or application to recognize plates as before or after meal	Medium
Manual input form	F5V1	Manual input form to record symptoms and mood via a 5-point Likert scale	High
Analytics	F6V1	Analytics to track user use of prototype	Medium
Opt-in or consent form	F7V1	Consent form for users during testing of prototype	High
Resident data storage	F8V1	Centralized storage of resident data so multiple devices can access resident data	High
Push notifications	F9V1	Notifications or prompts such as push notifications to attract the user back into the application	Medium

Next, Game Doctor produced a short video to guide participants on how to download the tool and use its initial features (ie, taking photos of plates of food before and after meals, logging symptoms via tick-box selections, and recording mood using a face emoji scale). The basic initial prototype was presented at the second meeting of AG 1. Game Doctor provided a walk-through of the basic prototype and presented different food recognition models Game Doctor. LogMeal was chosen as the food recognition model provider due to the size of its database and the accuracy of food recognition (refer to Table 3 for the models compared). This meeting lasted twice as long as the first session to allow for Game Doctor to talk through the functions of the tool and food recognition model, for the group to try it and test with singular foods, and then test with mixed foods over a restaurant meal. The prototype was tested by taking photos of fruits and vegetables after Game Doctor's presentation. Food in different presentations and textures was provided on plates for testing with the tool. Basic feedback was given when the group tested the tool with whole meals during dinner, with participants providing feedback on its usability. Other features identified to enhance the tool included having an "after" photo to calculate food eaten; adding an element to note whether food had been fortified (eg, high fat or additional protein added); including additional relevant medical conditions; and keeping the tool simple by, for example, using a tick-box approach and having symptom recording options appear when the "after" photo is uploaded.

Table 3. Different food recognition models available that could be integrated into the tool.

Provider	Pricing
Clarifai	Free for academic projects
LogMeal	Depends on project
FoodAI	To be confirmed
BiteAI	To be confirmed
Calorie Mama	To be confirmed

Key design issues noted in feedback included the positioning of the "confirm" screen, which sometimes overlapped the photo taken; how to enter a food if the database does not recognize it and provide a correct option; whether the highlighting circles link to each food on the plate; how to cancel the recognition system if there is an error in recognizing a food; and how plate size would be distinguished to calculate nutrition from different portion sizes. Important features discussed as being needed in the next build, based on feedback from using the tool, included (1) the ability to select resident ID and room number; (2) "before" and "after" photos; (3) adding a food fortification button to indicate the addition of butter, cream, protein powder, and so on; (4) having tick boxes for various symptoms that can be personalized to residents and structuring the page to follow the "after" photo or providing an option to enter symptoms later; (5) plate size option; (6) the option to add medical conditions to the user profile; and (7) having the symptoms page appear

right after the second photo, including options to indicate fullness, mood, swallowing difficulties, and so on.

Phase 2: Feasibility Testing of the Tool and Food Diaries

The next phase of the project involved further testing and feedback by users in their home settings. We recruited 10 community-dwelling older adults to test the tool for 2 weeks. Of the 10 participants, 9 (90%) managed to use the tool, while 1 (10%) had difficulty downloading it and withdrew from the study. Of the 9 participants, 8 (89%) used the tool for the full 2 weeks, while 1 (11%) used it for 1 week. Each day, they documented their food intake and any difficulties encountered with the tool in food diaries. In addition, members of AG 1 tested the tool in free-living conditions over 1 week and submitted comments about their experience through a brief web-based survey. The findings from the diaries and survey covered 3 topics: symptoms, features, and populations. Regarding symptoms, it was noted that the tool monitored a good range of symptoms specific to the target group; however, there were suggestions to include additional symptoms, such as thirst, confusion, and anxiety. Regarding features, it was reported that the tool was easy to use when it connected well with the food recognition model. Most of the group said that they did not feel that any key features were missing, although they noted that testing in care home settings would help identify site-specific needs. Suggested features included an "empty plate" button and a directory of foods. As for other populations and settings that might benefit from the tool, the survey responses highlighted community care, patients with dementia, family carers, and hospitals.

Phase 3

Overview

At the first meeting of AG 2, members suggested that care homes would face unique challenges in using the tool, different from those within the community. These might include potential difficulties in recognizing pureed food, fortified meals, or culturally specific foods (eg, haggis), residents with cognitive impairment not being able to use the tool, and the possibility for staff or visitors to assist with tool use. It was decided that we would trial the tool across a range of care homes with residents of varying abilities to gain direct feedback on these issues.

Care Home Testing and Focus Groups

The prototype was tested in 4 care homes with 8 staff members (n=2, 25% in each home) and in 1 care home with 6 residents (older adults: n=4, 67%; and younger adults [aged 54 y and 58 y] with learning disabilities: n=2, 33%). Of the 4 care homes, 3 (75%) did not recruit older adults for the testing due to the cognitive functioning status of their residents who would not be able to use the tool; therefore, the staff members used it themselves. The 6 residents who were recruited all tested the tool themselves. The staff recruited had from 6 months to 20 years of experience of working in the care sector. The older adults recruited were aged between 78 and 84 years and had lived in the care home for at least 6 months. After prototype testing in the care homes, the semistructured interviews with staff (4/8, 50%) and residents (1/6, 17%) identified (1) the pros and cons of using the tool in their care home, (2) how it can be adapted to enhance its usability, and (3) common symptoms reported in the tool. The resident who tested the tool was 1 (50%) of 2 younger residents with learning disabilities (aged 58 y). The main outcomes from these interviews are displayed in Textbox 1.

Textbox 1. Themes emerging from semistructured interviews with care home staff and residents.

Key design issues

- Wi-Fi issues slowed down the uploading of data
- Tool struggled to recognize pureed food
- Camera picked up patterns on plates as food
- Design features are needed to speed up the process

Key features

- Staff would prefer photo profiles of residents instead of names
- International Dysphagia Diet Standardization Initiative framework [29] and Malnutrition Universal Screening Tool [14] malnutrition risk score should be built in
- Traffic light system approach to identifying risk
- Self-choose whom to measure option
- Empty plate button
- Monitor fluid intake too
- Pureed food and fortification recognition is essential
- A notification or prompt system as a reminder

Positive feedback

- Recognized food well
- Residents liked monitoring their own food intake
- Enjoyed playing around with the tool and learned how to use it quickly

Key Design Issues and Usability

All care homes reported difficulties in using the tool when the home Wi-Fi connection was poor:

> *We had connection issues and found the app to be very slow or would freeze-could there be an easy offline system to use throughout the home that would automatically upload when a strong wifi connection was made?* [Staff, care home 2]

All care homes reported slow uploads, with the tool freezing; therefore, participants had difficulty clicking on the foods on the screen. This then led to slower identification of foods and a slow search for foods that were not identified, highlighting the need for features that would speed up this process.

A resident who was using the tool revealed that they did not have these issues because they just saved the photo to upload later when their Wi-Fi connection was stronger:

> *I found the tool easy to use, wifi is terrible here so I just waited until the signal was stronger and uploaded the photo then.* [Resident, care home 2]

Staff members expressed concern that if they did not upload photos and ratings immediately, they would not be able to do so later due to a change of shift or being given another role or task within the home. Furthermore, choking was identified as a major concern in all participating care homes, and many of the residents were on a pureed diet. The tool struggled to recognize pureed foods, highlighting the need for better pureed food recognition or an alternative solution:

> *A number of our residents are on pureed diets or we add thickening agents (which contain calories) to some of their foods, the tool could not properly detect it, although funnily enough we use molds shaped like the foods and it could sometimes pick that up.* [Staff, care home 1]

Staff members also wanted to add a measure of swallowing difficulties to the tool, which would link to food intake and the International Dysphagia Diet Standardization Initiative (IDDSI) framework [25]:

> *Having a feature and notes in IDDSI, malnutrition risk and fortification is really important for us and our residents.* [Staff, care home 4]

The current version did not link a swallowing difficulties feature, but adding this would encourage care homes to change the texture of the food.

Finally, when photos were taken of plates, the food recognition function of the tool sometimes misidentified the design patterns on the plate as food:

> *The plate was empty but it was saying the patterns were food and giving it a value.* [Staff, care home 3]

However, the more the tool was used, the less this became a problem because users worked out how to delete the circle around the patterned part that was misidentified as food. In addition, continued use improved the tool's accuracy through machine learning, reducing misidentifications over time.

Key Features Missing From the Tool

Regarding features needed in the tool for the care home setting, all homes reported the need for a resident profile instead of the current name function, which takes users straight to the diary:

> *We need a profile with the resident's picture which we choose from to quickly find the resident—we have a few residents with the same name.* [Staff, care home 2]

Staff members wanted to be able to upload a photo of each resident to click on instead of selecting a name, which would take them directly to the resident's profile. This profile would incorporate relevant information about the resident such as weight, changes in weight, issues with previous foods such as stomach problems or swallowing difficulties, risky foods and allergies, IDDSI [29] and MUST [14] scores, medications that may affect diet, and a category for food monitoring (high risk, medium risk, or low risk). Having the malnutrition risk score assessed via the tool and keeping track of IDDSI scores was a high priority for staff. These features would support staff, particularly new staff who are less familiar with a resident, in ensuring that the correct food is provided:

> *Choosing the picture would help with our new staff or our bank staff to recognize the resident, we have quite a high turnover here so a photo would quicken the recognition process.* [Staff, care home 4]

The profile would be accessed, leading to the diary, and a warning signal would be produced when a picture is taken to highlight risky foods for that resident. This was suggested as an important future feature:

> *If the tool was able to actually warn or prompt staff of certain foods it would really support our staff-we have a lot of residents here who have allergies or are choking hazards so it would support staff unfamiliar with these residents when giving them their meals.* [Staff, care home 2]

A key concern for care home staff using the tool was the time it would require to take the photo and deliver a meal to the resident while it was still hot. Some homes had 80 residents, and staff would not be able to monitor everyone in addition to their other duties. Upon further discussion, a traffic light system was suggested:

> *Having something like a traffic light system for overall nutrition risk of the resident as well as risk of the food would be brilliant. We could then monitor our resident over time and give more detailed reports when the dietitian came to who was priority.* [Staff, care home 3]

This would be a system that would categorize individuals as *green* (does not need regular monitoring), *amber* (needs monitoring but is not of major concern), and *red* (needs monitoring and may need observation or support with eating due to key issues such as swallowing difficulties or choking concerns). Care staff also wanted to be able to choose which residents to monitor because this would help reduce the time spent using the tool by prioritizing residents who needed monitoring the most while eliminating those who did do not require it.

Given concerns over staff time, methods to speed up the tool's use were highlighted as essential components to integrate. One suggestion was allowing staff to self-select individuals they needed to monitor based on their own experience of who needed support the most:

> *We know our staff well, we know who has problems when it comes to eating, and we know who to highlight to our manager, if we could choose who to monitor and have the tool to support our statements it would help.* [Staff, care home 1]

Another suggestion was to add an "empty plate" button:

> *Being able to just even say they ate everything would really help.* [Staff, care home 4]

This would eliminate the need for a second photo. A "half-finished" button was also suggested; however, it was highlighted that this could lead to users not knowing what parts of the meal had been eaten, resulting in an inaccurate estimate of nutritional intake:

> *Being able to click that they ate half would quicken things up, but saying that, we have certain residents who only eat their meat and leave their veggies.* [Staff, care home 2]

Having a database with preexisting photos based on current care home menus was identified to speed up use of the tool, with the staff only needing to take an "after" photo. Another suggestion to expedite the process was to work with the care homes to integrate their menus into the tool. Of the 4 care homes, 3 (75%) reported a 3-week menu that could be incorporated, and staff would pick from preassigned photos of each meal rather than take "before" photos of all meals. They would then only need to capture "before" photos for individuals with smaller portions or meal modifications (eg, substitution or fortification). Afterward, staff would take "after" photos of unfinished plates, using an "empty plate" button for completely finished meals.

Care staff mentioned the importance of monitoring fluid intake as well as food consumption. Linking fluid intake to bathroom breaks and urinary tract infection diagnoses would motivate staff to provide better fluid care for residents:

> *If we could monitor what they are drinking and then when we take them to the toilet, particularly during the night, it would be really useful so our staff can monitor who hasn't drank in a while. It would be even better if the tool could tell us when to stop giving them liquids or if certain liquids made them need more, to prevent them needing at night time.* [Staff, care home 3]

Staff felt that it would be useful to build a feature into the tool that gave personalized fluid feedback and indicated when best to give out fluids, the best type of fluids to give, and whether there should be any time restrictions to ensure the best level of hydration for each resident.

Given that most residents were on restricted diets—some requiring pureed food due to swallowing difficulties and others

needing fortified meals—having these dietary modifications recognized by, or incorporated into, the tool was deemed essential. Staff identified taking multiple photos as the only feasible solution that would still provide accurate nutritional information. This process would involve taking (1) a photo before pureeing, (2) a photo of the pureed meal, and (3) a photo after consumption. At the second meeting of AG 2, it was confirmed that further discussions with app developers would be needed to determine the quickest and most accurate method for implementing this approach.

Residents who tested the tool suggested having a notification or prompt system to remind them to use it. They mentioned that they sometimes forgot to log their food intake and became upset when they missed the opportunity to do so:

> *I eat at different times to the other residents, and was worried I wasn't allowed to use it out with mealtimes so I sometimes forgot to take pictures. This was annoying as I really enjoyed using the tool. Will I still have access to it?* [Resident, care home 2]

Residents reported that being able to set their own prompts or notifications would be beneficial. However, staff did not share this preference because they felt that they themselves would remember to use the tool at mealtimes as part of their work routine. The only instance in which staff considered a prompt system useful was for alerts warning them if a resident's meal contained an allergen or had previously posed a choking risk.

Care Home Staff Survey and Interviews

Altogether, 130 individuals working in the care industry completed the survey. Most of the respondents were health care staff working in a care home (62/130, 47.7%), followed by care workers (25/130, 19.2%), support staff (19/130, 14.6%), and managers (9/130, 6.9%). Of the 9 managers, 2 volunteered to participate in the postsurvey interview. Of the 130 respondents, 60 (46.2%) had worked in the care industry for >5 years, while 30 (23.1%) had worked for <2 years. When asked how they monitored residents' diets, most (114/130, 87.7%) did not provide specific examples or measurable approaches used in their care homes. Many mentioned menu planning, staff training, and nutritional analysis but without detailing the methods or explaining how food intake was tracked for each resident. Among those who provided examples, methods included written reports (9/16, 56%), diet diaries (4/16, 25%), and computer platforms (2/16, 13%). When asked about additional symptoms that they felt were important to monitor in residents but were not available in the current version of the tool, the most commonly mentioned were hunger (45/130, 34.6%), mood (41/130, 31.5%), and swallowing difficulties (39/130, 30%). However, these symptoms were already included in the tool but may need to be more visually prominent and not limited to being reported only after meals. Other suggested symptoms included tiredness and pain. When asked which features they felt should be added that were not available in the current version, the most common responses were profiles that use photos to identify resident rather than just their name (42/130, 32.3%), a traffic light system (36/130, 27.7%), and integration of IDDSI and MUST scores (36/130, 27.7%). Additional suggestions aligned with feedback from individual care homes, including a regular weigh-in section and fluid intake monitoring. These suggestions were discussed at the second meeting with AG 2, focusing on developing a basic version of the tool that meets general needs across care homes, with the option for individual care homes to add personalized features if deemed essential, provided they were willing to cover the additional cost.

The survey data were supported by the interviews with care home managers. Both managers reported using multiple methods to monitor residents' diets, although these were not always consistent across staff members or health professionals. They noted that having an app capable of tracking diet, identifying malnutrition, and monitoring associated symptoms would help standardize care across residents and life situations. The managers agreed that such a tool would allow dietary monitoring beyond the care home, including when residents were out with family or hospitalized, ensuring that symptoms could be monitored during these periods. A traffic light system that was visible to staff but not residents was considered highly important to ensure close monitoring of only residents who needed it. The managers were enthused about the tool and using it as part of care but had some concerns. They were worried about the time required to take photos and monitor symptoms, but when prompted during the interview, they came up with strategies themselves to speed up the process. Both managers expressed concerns about the cost of the tool and its uptake. While they acknowledged the tool's value, they noted that traditional paper methods were more economical unless the tool provided clear value for money. When discussing ways to make this feasible, they suggested (1) making it free for National Health Service (NHS) users, (2) implementing a tiered fee structure for access to different tool features, (3) generating revenue from external users to help subsidize the costs for care homes, and (4) finding funding to support its implementation in care homes.

Discussion

Principal Findings

We co-designed a prototype tool that enabled users to take photos of food plates using Android phones or tablets and link foods to a variety of chosen symptoms and mood ratings. The prototype was then tested in care homes. While the ultimate version of the tool would also calculate nutritional intake from the photos, the aim of this study was to co-design the tool to maximize usability. The co-design involved older adults, care home staff, dietitians, app developers, and researchers. The repeated design and test approach ensured that the prototype was co-designed with AGs, with participants' feedback used to make incremental improvements before testing whether the suggested features were usable and integral to the overall objectives of the tool. This agile methodology implemented by Game Doctor allowed for rapid prototyping and testing of each of the 3 builds by end users. Testing in the community and later in care homes revealed what worked well and what needed updating as well as features that should be included to make the tool specifically suitable for use in care home environments. This collaboration with end users was vital to developing a tool suited for diet and symptom monitoring in care homes.

Care home staff and residents provided feedback on how to ensure that the tool could be implemented in their care home and were very positive about using it overall, although it should be acknowledged that this may have been influenced by the relative experience of the staff. However, the range of experience across the testers gives us confidence that we have captured a range of views. All participants emphasized the tool's value in supporting staff; facilitating quicker diagnosis of dietary-related symptoms; and enabling earlier detection of malnutrition, which is an important issue in care homes [7]. They reported that they would not need to monitor every resident or use every aspect of the tool but that it would be very valuable to use and could additionally form part of their marketing approach to attract residents.

Care home staff and residents suggested several important features and ways to increase usability. Priority features that will be integrated into future design are enhanced resident profiles with a traffic light system, integration of IDDSI [29] and MUST [14] scores, and recognition of pureed food and fortified meals. Traffic light systems have been found to be successful in dietary monitoring previously by encouraging users to choose lower-calorie alternatives [30]. A similar approach could be applied to identifying "safe foods" in the current targeted group. This would circumvent the need for diet diaries and bringing in a dietitian to interpret data [15], potentially accelerating the recognition of malnutrition with the addition of a nutrition calculator and other tool features to enable appropriate actions to be taken before hospitalization becomes necessary. Previous tools [22,23] that have used photo-based dietary assessment have found this to be a valid, feasible, and user-friendly approach to monitoring dietary intake. However, these tools have focused on younger populations, who may have different dietary requirements and greater familiarity with technology compared to the current target group.

Regarding wider use across care homes and potential methods of commercialization, the care staff survey and manager interviews indicated a preference for the tool to be free for NHS users. This would support its adoption across services and ensure that those who most needed it could fully benefit from its features. Finding funding from the local council or government or having the cost subsidized by nonresident users was brought up as another way to keep costs low or eliminate them for care homes. Managers suggested that even staff members interested in using the tool for their own dietary monitoring could help subsidize costs. When discussing payment for the tool, managers remained positive about using it, provided it was not too expensive, which would divert funding from other activities. Their preference was to pay an annual fee that would give them access to the tool, including updates and maintenance support. A tiered fee structure was also suggested, where costs would increase based on the number of features accessed. Although the managers would prefer the tool to be free, they were still keen to use it even if it was associated with a small fee. The global revenue for nutrition apps was projected to reach US $5.4 billion in 2024, with an expected annual growth rate of 11.2% [31]. This highlights the popularity of dietary monitoring and the importance of working with care home organizations to develop a financially feasible model for the tool.

Co-design between end users, the tool developers, and the research team was integral to this tool's production. As previous tools [22,23] and reviews of image-assisted dietary assessment methods [32] have focused on younger populations, it was essential to involve older adults in the codevelopment process. The feedback gathered played a crucial role in shaping the tool's functionality and features, ensuring that it was suited to the intended environment and users. This emphasizes the importance of co-design and coproduction in this type of research and innovation [27]. Applying the 3-sprint approach to design enabled a continual process of refinement, resulting in the production of a tool created by and for older adults and staff in care homes. Without this iterative process, the research team could have estimated which features might be important for users in care homes, but direct lived-experience evidence was essential to verifying whether these were the right features to incorporate and ensuring the tool's feasibility for the intended end users. However, despite this inclusive approach to ensuring that the tool incorporated the most important features for the setting, several technical improvements are still needed, along with the integration of new features; for example, because this study focused on co-design and feasibility testing, it is not yet clear how well the tool scores in terms of the accuracy of food recognition, and further testing, as well as comparison to similar tools, will be needed to ensure that this critical aspect performs well. Future versions would need to use a specially trained model that is linked to known menus; this should improve food recognition substantially. Other tools have reported mixed accuracy estimates, ranging from 9% to 63% [33]. One advantage of our tool is that it allows manual corrections or direct entry of specific foods and can be directly linked to a care home's menu plan to increase the accuracy of food recognition. Many similar tools have been criticized for their inability to accurately assess the quantity of food eaten [33]. Our tool addresses this limitation by incorporating "before" and "after" meal photos, enabling it to calculate the actual amount consumed rather than just identifying the food on the plate. Future research is needed to evaluate food quantity estimation as well as food recognition performance. However, the tool has the potential to simplify and speed up malnutrition recognition and dietary monitoring in later life among older adults, who are at higher risk than younger populations.

Limitations

This study is not without limitations. While the use of co-design and coproduction ensured that the tool included relevant features suited to its intended setting, our testing sample lacked diversity. All care homes recruited were private rather than council funded, and all AG members and tool testers were of White British origin. However, this reflects the typical ethnic composition of the Scottish population [34], and there was diversity in terms of gender, age, occupation, and specific care home site across participants. Future research would benefit from recruiting more diverse samples, including participants from both council-funded and private care homes, as well as recruiting across a range of geographic locations, including inner-city areas where communities are more ethnically diverse. A second limitation was the specific database used, which was developed in Spain and optimized for a more European-style diet. As a result, it did

not include many foods commonly consumed in Scotland and misidentified some foods as typical Spanish dishes (eg, paella). However, a future direction for this research is to create a bespoke database of food consumed in Scottish care homes as well as exploring the possibilities of working with individual homes' menu planners to create a tool specific to each home. This approach could be fundamental to the economic model for maintaining and continually updating the tool as a way of charging for a bespoke service rather than tool *use* in future, allowing us to provide the tool free in other settings, such as the NHS. One limitation of the co-design and coproduction methodology was the diversity of feedback regarding the tool and its usability and key features. While some care homes wanted enhanced features, such as malnutrition risk scores, a traffic light system, and fluid measurement with urinary continence tracking, others preferred to keep the tool's functions simple to enable ease of use. Further discussions clarified that the enhanced features could be optional, tailored to individual care homes' needs, and potentially cost associated, which eliminated the contradictions in the feedback.

Future Directions

Building on the outlined future directions, the next steps for this research project are to incorporate the care home feedback into the next build of the tool to ensure its suitability for care homes. To do this, we plan to apply steps 3 to 6 of the 6SQuID framework [24], test the tool in a more diverse sample of care homes, link it to a specially developed database of food photos reflecting the typical Scottish care home diet, and explore cost-effectiveness options and sustainability strategies to ensure that the tool remains up to date.

Conclusions

This journey through the co-design of a digital food recognition tool has revealed both its considerable potential and areas for further design improvement and feature enhancement. As we progress toward further prototype refinement and wider testing in diverse care home settings, the feedback and recommendations from this project will guide our approach, ensuring that the tool is practical, feasible, and robustly designed to comprehensively monitor diet and identify malnutrition risk and food-related problematic symptoms in older adults in residential care homes.

Acknowledgments

The authors acknowledge the artificial intelligence company specializing in food recognition and nutrition analysis LogMeal (AIGecko Technologies SL, Artificial Intelligence and Deep Learning Food Division) for the use of its database and thank all advisory group members and tester participants. This research was funded by a UK Research and Innovation Healthy Ageing Catalyst grant (ES/X006271/1), an Economic and Social Research Council–Stirling Social Science Impact Acceleration Account institutional block award to the University of Stirling, and a specific exploratory–Impact Acceleration Account and routes to commercialization grant to JC and ACW (STI-23/24-P0001). The article processing charge was funded by the University of Stirling. The funders had no role in the study's design; the collection, analyses, or interpretation of data; the writing of the manuscript; or the decision to publish the results.

Conflicts of Interest

None declared.

References

1. Stratton R, Cawood A, Anderson L, Burden S, Collery R, Holmes E, et al. Malnutrion and nutrional care survey in adults: UK malnutrion awareness week, October 2022. BAPEN. 2022. URL: https://www.bapen.org.uk/pdfs/reports/mag/national-survey-of-malnutrition-and-nutritional-care-2022.pdf [accessed 2025-02-06]
2. Malnutrition. World Health Organization. URL: https://www.who.int/health-topics/malnutrition#tab=tab_1 [accessed 2024-12-13]
3. Roe A. The importance of eating well in later life. Age UK. 2018 Sep 30. URL: https://www.ageuk.org.uk/discover/2018/importance-of-eating-well-in-later-life/ [accessed 2023-05-12]
4. Profile of the older population living in England and Wales in 2021 and changes since 2011. Office for National Statistics. 2023 Apr 3. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationlivinginenglandandwalesin2021andchangessince2011/2023-04-03#:~:text=2.-,Population%20ageing,from%2016.4%25%20to%2018.6%25 [accessed 2024-03-26]
5. Dent E, Wright OR, Woo J, Hoogendijk EO. Malnutrition in older adults. The Lancet 2023 Mar;401(10380):951-966. [doi: 10.1016/s0140-6736(22)02612-5]
6. Norman K, Haß U, Pirlich M. Malnutrition in older adults-recent advances and remaining challenges. Nutrients 2021 Aug 12;13(8):2764 [FREE Full text] [doi: 10.3390/nu13082764] [Medline: 34444924]
7. State of the nation 2021: older people and malnutrition in the UK today. The Malnutrition Task Force. 2021. URL: https://www.malnutritiontaskforce.org.uk/sites/default/files/2021-10/State%20of%20the%20Nation%202020%20F%20revise.pdf [accessed 2023-03-26]
8. Schenker S. Undernutrition in the UK. Nutr Bull 2003 Mar 17;28(1):87-120. [doi: 10.1046/j.1467-3010.2003.00303.x]

9. Stratton R, Smith T, Gabe S. Managing malnutrition to improve lives and save money. BAPEN. 2018 Oct. URL: https://www.bapen.org.uk/pdfs/reports/mag/managing-malnutrition.pdf [accessed 2025-02-06]
10. Fávaro-Moreira NC, Krausch-Hofmann S, Matthys C, Vereecken C, Vanhauwaert E, Declercq A, et al. Risk factors for malnutrition in older adults: a systematic review of the literature based on longitudinal data. Adv Nutr 2016 May;7(3):507-522 [FREE Full text] [doi: 10.3945/an.115.011254] [Medline: 27184278]
11. Wei K, Nyunt MS, Gao Q, Wee SL, Ng TP. Frailty and malnutrition: related and distinct syndrome prevalence and association among community-dwelling older adults: Singapore longitudinal ageing studies. J Am Med Dir Assoc 2017 Dec 01;18(12):1019-1028. [doi: 10.1016/j.jamda.2017.06.017] [Medline: 28804010]
12. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 Feb;27(1):5-15. [doi: 10.1016/j.clnu.2007.10.007] [Medline: 18061312]
13. Dean M, Raats MM, Grunert KG, Lumbers M, Food in Later Life Team. Factors influencing eating a varied diet in old age. Public Health Nutr 2009 Dec;12(12):2421-2427. [doi: 10.1017/S1368980009005448] [Medline: 19344544]
14. Introducing 'MUST'. BAPEN. URL: https://www.bapen.org.uk/must-and-self-screening/introducing-must/ [accessed 2024-01-29]
15. Ortega RM, Pérez-Rodrigo C, López-Sobaler AM. Dietary assessment methods: dietary records. Nutr Hosp 2015 Feb 26;31 Suppl 3:38-45 [FREE Full text] [doi: 10.3305/nh.2015.31.sup3.8749] [Medline: 25719769]
16. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
17. Fakhrou A, Kunhoth J, Al Maadeed S. Smartphone-based food recognition system using multiple deep CNN models. Multimed Tools Appl 2021 Aug 12;80(21-23):33011-33032. [doi: 10.1007/s11042-021-11329-6]
18. Sahoo D, Hao W, Xiongwei W, Le H, Achananuparp P, Lim EP, et al. FoodAI: food image recognition via deep learning for smart food logging. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019 Presented at: KDD '19; August 4-8, 2019; Anchorage, AK URL: https://arxiv.org/abs/1909.11946
19. Banerjee S, Mondal AC. Nutrient food prediction through deep learning. In: Proceedings of the 2021 Asian Conference on Innovation in Technology. 2021 Presented at: ASIANCON 2021; August 27-29, 2021; Pune, India. [doi: 10.1109/asiancon51346.2021.9545014]
20. Zeng J, Zhang J, Li Z, Li T, Li G. Prediction model of artificial neural network for the risk of hyperuricemia incorporating dietary risk factors in a Chinese adult study. Food Nutr Res 2020;64 [FREE Full text] [doi: 10.29219/fnr.v64.3712] [Medline: 32047420]
21. Ho DK, Tseng SH, Wu MC, Shih CK, Atika AP, Chen YC, et al. Validity of image-based dietary assessment methods: a systematic review and meta-analysis. Clin Nutr 2020 Oct;39(10):2945-2959. [doi: 10.1016/j.clnu.2020.08.002] [Medline: 32839035]
22. Prinz N, Bohn B, Kern A, Püngel D, Pollatos O, Holl RW. Feasibility and relative validity of a digital photo-based dietary assessment: results from the Nutris-Phone study. Public Health Nutr 2018 Mar 6;22(7):1160-1167. [doi: 10.1017/s1368980018000344]
23. Budiningsari D, Syahrian F. Validity of a digital photo-based dietary assessment tool: development and initial evaluation. Nutr Health 2024 Jul 23:2601060241239095. [doi: 10.1177/02601060241239095] [Medline: 39043223]
24. Wight D, Wimbush E, Jepson R, Doi L. Six steps in quality intervention development (6SQuID). J Epidemiol Community Health 2016 May;70(5):520-525 [FREE Full text] [doi: 10.1136/jech-2015-205952] [Medline: 26573236]
25. Oliver K, Kothari A, Mays N. The dark side of coproduction: do the costs outweigh the benefits for health research? Health Res Policy Syst 2019 Mar 28;17(1):33 [FREE Full text] [doi: 10.1186/s12961-019-0432-3] [Medline: 30922339]
26. NIHR guidance on co-producing a research project. National Institute for Health and Care Research. 2024. URL: https://www.learningforinvolvement.org.uk/content/resource/nihr-guidance-on-co-producing-a-research-project/? [accessed 2024-02-23]
27. Smith B, Williams O, Bone L, the Moving Social Work Co-production Collective. Co-production: a resource to guide co-producing research in the sport, exercise, and health sciences. Qual Res Sport Exerc Health 2022 Mar 19;15(2):159-187. [doi: 10.1080/2159676x.2022.2052946]
28. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec 04;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
29. The IDDSI framework (the standard). International Dysphagia Diet Standardisation Initiative. URL: https://www.iddsi.org/standards/framework [accessed 2025-02-25]
30. Osman M, Thornton K. Traffic light labelling of meals to promote sustainable consumption and healthy eating. Appetite 2019 Jul 01;138:60-71. [doi: 10.1016/j.appet.2019.03.015] [Medline: 30880087]
31. Nutrition apps- worldwide. Statista. URL: https://www.statista.com/outlook/hmo/digital-health/digital-fitness-well-being/health-wellness-coaching/nutrition-apps/worldwide [accessed 2024-12-13]
32. Höchsmann C, Martin CK. Review of the validity and feasibility of image-assisted methods for dietary assessment. Int J Obes (Lond) 2020 Dec 08;44(12):2358-2371 [FREE Full text] [doi: 10.1038/s41366-020-00693-2] [Medline: 33033394]

33. Van Asbroeck S, Matthys C. Use of different food image recognition platforms in dietary assessment: comparison study. JMIR Form Res 2020 Dec 07;4(12):e15602 [FREE Full text] [doi: 10.2196/15602] [Medline: 33284118]
34. Ethnicity. Scotland's Census. URL: https://www.scotlandscensus.gov.uk/census-results/at-a-glance/ethnicity/ [accessed 2025-02-25]

Abbreviations

6SQuID: 6 steps for quality intervention development
IDDSI: International Dysphagia Diet Standardization Initiative
MUST: Malnutrition Universal Screening Tool
NHS: National Health Service

Edited by J Wang; submitted 23.07.24; peer-reviewed by A Coltman, O Ueland, DR Crabtree; comments to author 11.11.24; revised version received 18.12.24; accepted 30.01.25; published 03.03.25.

Please cite as:
Connelly J, Swingler K, Rodriguez-Sanchez N, Whittaker AC
Identifying Food Preferences and Malnutrition in Older Adults in Care Homes: Co-Design Study of a Digital Nutrition Assessment Tool
JMIR Aging 2025;8:e64661
URL: https://aging.jmir.org/2025/1/e64661
doi:10.2196/64661
PMID:40053797

Review

Values of Stakeholders Involved in Applying Surveillance Technology for People With Dementia in Nursing Homes: Scoping Review

Daniëlle van Gaans-Riteco[1,2], MSc; Annerieke Stoop[1], PhD; Eveline Wouters[1,3], PhD, MD

[1]Academic Collaborative Center Care for Older Adults, Tranzo, Scientific Center for Care and Wellbeing, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands

[2]Groenhuysen, Roosendaal, The Netherlands

[3]School for Allied Health Professions, Fontys University of Applied Science, Eindhoven, The Netherlands

Corresponding Author:
Daniëlle van Gaans-Riteco, MSc
Academic Collaborative Center Care for Older Adults
Tranzo, Scientific Center for Care and Wellbeing
Tilburg School of Social and Behavioral Sciences, Tilburg University
Professor Cobbenhagenlaan 125
Tilburg, 5037DB
The Netherlands
Phone: 31 134662969
Email: d.p.c.vangaans-riteco@tilburguniversity.edu

Abstract

Background: Due to the progressive nature of dementia, concerns about the safety of nursing home residents are frequently raised. Surveillance technology, enabling visual and auditory monitoring, is often seen as a solution for ensuring safe and efficient care. However, tailoring surveillance technology to individual needs is challenging due to the complex and dynamic care environment involving multiple formal and informal stakeholders, each with unique perspectives.

Objective: This study aims to explore the scientific literature on the perspectives and values of stakeholders involved in applying surveillance technology for people with dementia in nursing homes.

Methods: We conducted a scoping review and systematically searched 5 scientific databases. We identified 31 articles published between 2005 and 2024. Stakeholder characteristics were extracted and synthesized according to the theory of basic human values by Schwartz.

Results: In total, 12 stakeholder groups were identified, with nursing staff, residents, and informal caregivers being the most frequently mentioned. Among stakeholder groups close to residents, values related to benevolence, security, conformity, and tradition were most commonly addressed. Furthermore, values such as self-direction, power, and achievement seemed important to most stakeholder groups.

Conclusions: Several stakeholder groups emphasized the importance of being and feeling involved in the application of surveillance technologies. In addition, they acknowledged the necessity of paying attention to stakeholders' perspectives and values. Across these stakeholder groups, values related to benevolence, security, and self-direction were represented, although various stakeholders assigned different meanings to these values. Awareness of stakeholders' perspectives demands a willingness to acknowledge each other's values and bridge differences.

(JMIR Aging 2025;8:e64074) doi:10.2196/64074

KEYWORDS

surveillance technology; nursing home; stakeholders; values; dementia; safety

Introduction

Background

Globally, people are living longer. Every country in the world expects growth in the number and proportion of older persons [1]. As the population ages, diseases such as dementia are diagnosed more often because age is their strongest known risk factor [2,3]. Dementia is a major cause of disability and dependency, affecting cognitive abilities and behavior, leading to an inability to signal when help is needed, which is associated with safety concerns [3,4]. Compared with 2015, the number of people affected with dementia will triple by 2050 [3].

In several countries worldwide, the number of care professionals is insufficient to meet the growing care demands [5]. In addition, global shortages of skilled care professionals will increase due to the aging workforce [6,7]. To meet the increasing care demands, documents from the Netherlands government show that by 2024, 1 in 4 employees should work in the health care sector, a figure that should rise to 1 in 3 by 2060 [8,9]. One of the proposed solutions to this challenge is allowing people with dementia to live at home as long as possible. However, due to their increasing care demands, a need for long-term care (LTC) facilities providing high-quality intensive dementia care will continue. Studies in different countries worldwide indicate that approximately 30% to 40% of people with dementia will eventually need a care environment in nursing homes [10-12]. The Netherlands is known for its high percentage of residents receiving end-of-life care in nursing homes, which makes nursing homes the most frequent place of death [13]. Consequently, there has been a large increase in health care expenditures for the population with dementia, especially in LTC facilities [12].

One recommendation to address these societal challenges in home care and nursing homes is to foster investment in health technologies that contribute to sustainable and high-quality care for people with dementia, such as assistive and innovative care technologies [3,14,15]. Care technologies can delay or replace admission in a nursing home and reduce the workload of nursing staff and informal caregivers in community care and nursing homes [16,17]. There are different types of care technologies. One is surveillance technology, which allows visual and auditory monitoring and registration of events, including residents' activities. Surveillance technologies include tagging and tracking technology, sensors, and audio and video surveillance [18,19]. Surveillance technologies are increasingly focused on supporting autonomy and respecting privacy while enhancing safety and individualized care for people with dementia [7,20-23].

Surveillance technologies are often regarded as a solution for ensuring safe and efficient health care, including in nursing homes [14,19,20]. These technologies have the potential to provide high-quality care and relieve nursing staff as staff shortages increase [7]. Due to the potential benefits of using surveillance technologies for quality of life and care, general attitudes toward these technologies have become more positive [16]. Nonetheless, surveillance technologies can affect privacy, autonomy, and freedom of movement [16,20,21]. Therefore, the use of these technologies should comply with regulations governing privacy and involuntary care, including requirements for subsidiarity, proportionality, and expediency. In addition, the use of surveillance technologies has to be justified in the care plan [24-26].

The application of surveillance technologies for people with dementia living in nursing homes is complex in practice. Successfully implementing care technologies, including surveillance technologies in psychogeriatric nursing homes, appears to be challenging as it involves more than just implementing a technological application successfully used elsewhere [27]. One of the greatest challenges in implementing care technologies such as surveillance technology seems to be integrating technology into the care process. Surveillance technologies affect residents and other primary stakeholders, such as residents' representatives and formal caregivers [16,28]. The involvement of these primary stakeholders, also known as end users, and secondary stakeholders such as managers, information and communication technology (ICT) employees, developers, and vendors of surveillance technologies is necessary to increase stakeholder commitment [28].

Early involvement of relevant stakeholders increases the likelihood of successful implementation [29]. A prerequisite for their involvement is knowing and acknowledging stakeholders' cultures, perspectives, and interests [29]. Given the broad spectrum of stakeholders involved in applying surveillance technologies for people with dementia in nursing homes, there is a great diversity of backgrounds, resulting in differences in values and interests [27,30]. Values represent what is (most) important to people and direct their attitudes, behaviors, and actions [31,32]. Differences in stakeholder values and interests can complicate the creation of support among stakeholders [30]. In addition, dealing with different perceptions and values among a range of stakeholders is a major challenge, further exacerbated by a limited understanding of stakeholders' values [33]. Therefore, a knowledge of these values can help explain decision-making processes, attitudes, and behaviors of persons or groups in different contexts [31,34,35].

This situation necessitates exploring the stakeholder groups that are involved in implementing surveillance technologies, and their respective values. Nursing staff and informal caregivers' attitudes toward using surveillance technologies [18,36] and their ethical dilemmas when using surveillance technology in psychogeriatric nursing homes have been explored [23,36-38]. For example, Rostad and Stokke [39] noted the high complexity of the LTC setting, involving "wicked problems," such as many and changing stakeholders, competing interests, and disagreements regarding the nature of problems. However, it remains unclear what these competing interests consist of. In addition, little is known about the variation in the perspectives and values of the stakeholders.

This Study

To the best of our knowledge, no literature review has been conducted to explore the stakeholders involved, their perspectives, and values in the application of surveillance technologies for people with dementia living in nursing homes. Therefore, this scoping review aimed to explore which stakeholders are described in the scientific literature concerning

surveillance technologies for people with dementia in nursing homes. In addition, we seek to identify what is known about these stakeholders' values.

Methods

Overview

We conducted a scoping review to systematically explore, map, and synthesize the characteristics of stakeholders involved in applying surveillance technology for people with dementia in nursing homes and identify existing knowledge gaps. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist was used as a guideline for this review [40] (Multimedia Appendix 1). The corresponding steps were followed [41]: (1) identifying research questions; (2) identifying relevant literature in databases; (3) selecting the literature; (4) charting the data; and (5) collating, summarizing, and reporting the results.

Identifying the Research Questions

The research questions formulated were as follows: (1) Which stakeholders are involved in applying surveillance technology for people with dementia residing in psychogeriatric wards in nursing homes? and (2) What is known about the values of these stakeholders?

Identifying Relevant Literature

We believed that articles of interest had been published in psychological, health care, medical, nursing, and technological journals. Therefore, we conducted a literature search using the following databases: MEDLINE, CINAHL, PsycINFO, ACM Digital Library, and IEEE Xplore. Search terms encompassed the LTC setting and the use of surveillance technology for people with dementia. A search string for each database was developed and programmed with the help of an information specialist (Multimedia Appendix 2). A search was first performed in April 2022 and fully updated in August 2023 and December 2024.

Studies were eligible for inclusion if surveillance technology was applied to people with dementia in nursing homes or a comparable 24×7 LTC setting. The use of surveillance technology was evaluated or monitored using qualitative, quantitative, or mixed methods designs. Furthermore, studies had to be peer reviewed, written in English or Dutch, and published between 2002 and 2023. The year 2002 was chosen because it was then that literature on the implementation of surveillance technologies became increasingly prevalent. In addition, studies had to mention the stakeholders who were involved in the process of applying surveillance technology. Studies mostly focusing on assistive technologies, such as automatic lights, or supportive technologies, for example, medication dispensers, health care apps for managing chronic diseases, etc were excluded. Studies conducted in an experimental or laboratory setting and nonoriginal research, such as scoping reviews and systematic reviews, were excluded.

Literature Selection

First, duplicate studies were removed. Title screening was performed by one of the authors (DvG-R), and in case of doubt, one of the other authors (AS) was consulted. Two authors (DvG-R and AS) independently screened abstracts using the literature review management tool Rayyan (Rayyan Systems Inc) [42]. The full text of articles considered eligible by both authors was reviewed for relevance. In all the selection steps, the results were compared and discussed until a consensus was reached. In case of doubt, the third author (EW) was consulted.

Charting the Data

A format for further data extraction was agreed upon and included the title, authors, year, country, aim, study design, method of data collection, study population, sample size, setting, technology type, an overview of the results per identified stakeholder, and limitations for this scoping review. Using this format, 2 authors (DvG-R and AS) independently reviewed 10 (32%) of the 31 included articles. When the reviews were compared, only minor differences were found. The remaining articles were reviewed by DvG-R, who consulted one of the two other authors (AS or EW) when appropriate. When no consensus about data extraction was reached, the other author (EW or AS) was consulted.

Collating, Summarizing, and Reporting the Results

We categorized the findings from each article per stakeholder group in data extraction forms. Subsequently, an overview of results per stakeholder group was compiled. Through an inductive process, we categorized our findings into frequently mentioned words, such as acceptance; privacy; safety; freedom of movement; person-centered care; quality of life; quality-of-care characteristics; technology characteristics; and resident characteristics, involvement, concerns, and their values. The findings were linked to human values to deepen an understanding of stakeholders' perspectives. Describing and defining values is considered complex, and analyzing them is an even greater challenge [43]. Therefore, we used the theory of basic human values by Schwartz et al [31], an empirically tested framework of values that is recognized across many cultures [44]. This theory is an important and well-known theory and is widely used to predict attitudes and behaviors in different contexts and situations [34,35]. These values are grounded in the 3 universal requirements of human existence: the needs of individuals as biological organisms, requirements of coordinated social interaction, and survival and welfare needs of groups [31]. This framework conceptualizes values ordered by importance relative to one another, and they form a system of priorities for groups, societies, and individuals [45]. The refined theory of basic human values has 19 values grouped into 4 higher-order categories as follows: openness to change, self-enhancement, conservation, and self-transcendence [31]. Table 1 presents the motivational goals of the Schwartz values based on the circular motivational continuum [31].

Table 1. The 19 values of Schwartz et al [31] explained in terms of their motivational goals based on the circular motivational continuum.

Higher order value and values	Conceptual definition in terms of motivational goals
Openness to changes	
Self-direction-thought	Freedom to cultivate one's own ideas and abilities
Self-direction-action	Freedom to determine one's own actions
Stimulation	Excitement, novelty, and change
Self-enhancement	
Hedonism	Pleasure and sensuous gratification
Achievement	Success according to social standards
Power-dominance	Power through exercising control over people
Power-resources	Power to control material and social resources
Conservation	
Face	Security and power through maintaining one's public image and avoiding humiliation
Security-personal	Safety in one's immediate environment
Security-societal	Safety and stability in the wider society
Tradition	Maintaining and preserving cultural, family, or religious traditions
Conformity-rules	Compliance with rules, laws, and formal obligations
Conformity-interpersonal	Avoidance of upsetting or harming other people
Self-transcendence	
Humility	Recognizing one's insignificance in the larger scheme of things
Benevolence-dependability	Being a reliable and trustworthy member of the ingroup
Benevolence-caring	Devotion to the welfare and well-being of ingroup members and being empathic
Universalism-concern	Commitment to equality, justice, and protection for all people
Universalism-nature	Preservation of the natural environment
Universalism-tolerance	Acceptance and understanding of those who are different from oneself

Results

General Findings

We identified 4116 unique studies. After screening titles, 298 (7.2%) abstracts were screened. After applying inclusion and exclusion criteria, we selected 91 (30.5%) articles for full-text screening. Eventually, we included 29 (32%) articles in this study. The search was updated in August 2023, and December 2024, resulting in 2 additional articles, bringing the total to 31 articles. The included studies were published between 2005 and 2023. Studies were conducted in 14 countries: The Netherlands (n=7, 23%), Finland (n=2, 6%), Denmark (n=1, 3%), Norway (n=2, 6%), Sweden (n=2, 6%), Spain (n=1, 3%), Germany (n=1, 3%), Switzerland (n=1, 3%), the United Kingdom (n=3, 10%), the United States (n=5, 16%), Canada (n=3, 10%), Mexico (n=1, 3%), Singapore (n=1, 3%), and China (n=1, 3%). A total of 10 (32%) articles had a quantitative research design, 18 (58%) articles had a qualitative design, and 3 (10%) articles had a mixed methods design. Figure 1 shows the flowchart of the selection process. The key characteristics of the included articles are available in Multimedia Appendix 3 [7,18-22,37,38,46-68].

In total, 12 stakeholder groups were identified in the publications. The most frequently identified stakeholder groups were nursing staff (23/31, 74%), residents (20/31, 65%), and informal caregivers (13/31, 42%). The stakeholder groups care managers (9/31, 29%), developers of surveillance technologies (7/31, 23%), physicians (5/31, 16%), LTC administrators (4/31, 13%), maintenance employees (2/31, 6%), ICT (2/31, 6%), vendors of surveillance technologies (1/31, 3%), project managers (1/31, 3%), and academics (1/31, 3%) were mentioned less frequently. A total of 9 (29%) of the 31 articles included 4 or more stakeholder groups in their research. Table 2 shows the stakeholders' frequency of occurrence in the included articles.

Figure 1. Flowchart of the selection process.

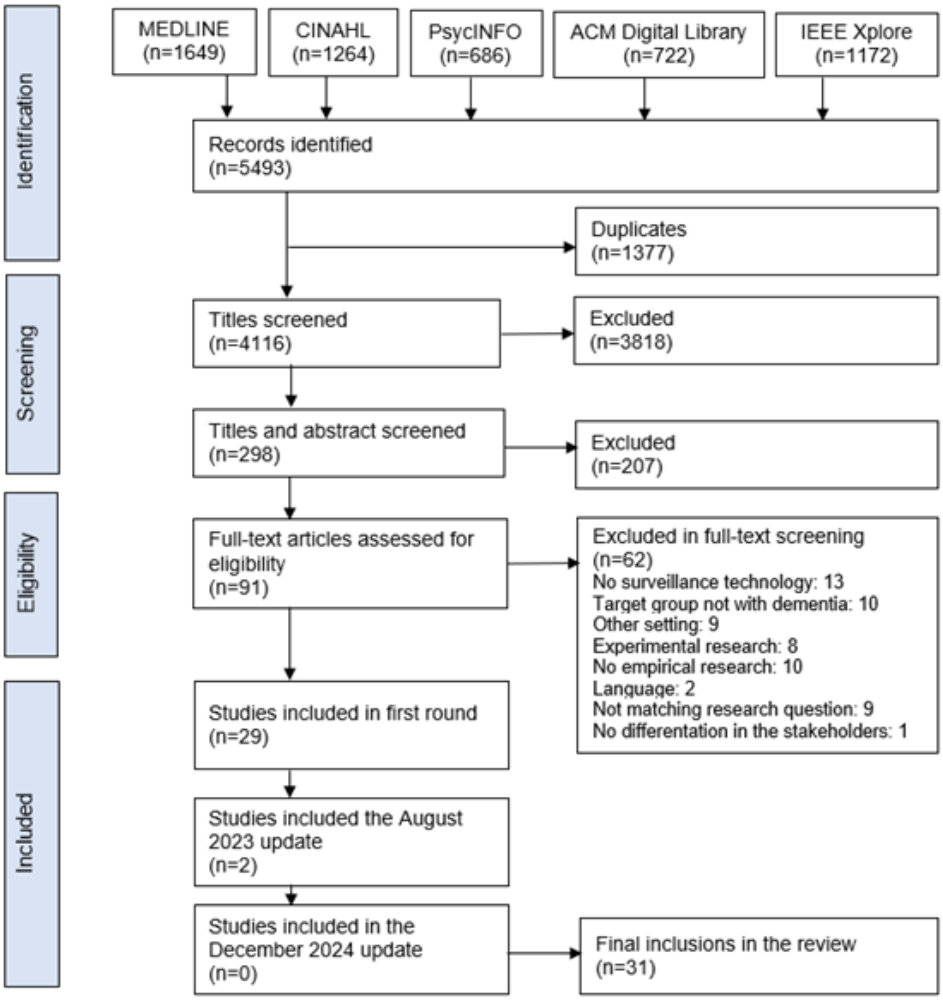


Table 2. Frequency of stakeholder occurrence in included articles and proxy perspectives[a].

Stakeholder	Frequency (N=31), n (%)	Proxy or partly proxy, n (%)[b]
Nursing staff	23 (74)	7 (30)
Residents	20 (65)	17 (85)
Informal caregivers	13 (42)	7 (53)
Managers	9 (29)	—[c]
Developers	7 (23)	—
Physicians	5 (16)	—
Long-term care administrators	4 (13)	—
Maintenance	2 (6)	—
ICT[d]	2 (6)	—
Vendors	1 (3)	—
Project managers	1 (3)	—
Academics	1 (3)	—

[a]Proxy perspective refers to the perspective of a stakeholder filled in by another stakeholder.

[b]Percentages in this column represent the proportion of articles in the corresponding cells in the "Frequency" column.

[c]Not applicable.

[d]ICT: information and communication technology.

Values of Stakeholders

Nursing Staff

Of all identified stakeholder groups, nursing staff (registered nurses, assistant nurses, and nurse aids) were mentioned in 23 (74%) out of 31 articles. Of these 23 articles, 7 (23%) articles were written from a proxy perspective of care managers (n=4, 17%), LTC administrators (n=2, 8%), and informal caregivers (n=1, 2%).

Several values of Schwartz were represented in the stakeholder group nursing staff. The values *benevolence* and *security* were strongly represented. The value *benevolence* is related to being helpful, supporting, and assisting those in need [31]. Staff's desire to be helpful and respond to the needs of residents also partly overlapped with the value *security*, which is characterized by striving to assure the safety of loved ones, taking precautions to prevent harm, and being warned in case of threats [45]. From the nursing staff perspective, surveillance technologies were most often applied to enhance the general safety of residents, mitigate the risk of falling, alert staff promptly, and contribute to efficient care [7,22,46]. Surveillance technologies can help nursing staff prioritize and direct their attention to where care and support are most needed [7,22,47]. Nursing staff expected monitoring technologies to increase the safety of residents [20], although they were aware of its limitations [21]. For example, they mentioned that surveillance technologies could not guarantee that residents would never fall [22]. Nursing staff rated residents' safety higher than they rated the experience of freedom [20], possibly due to the fear of being blamed for accidents or injuries to residents [22].

Because nursing staff felt responsible for their residents, the former expressed hesitation toward entirely relying on new surveillance technologies [48]. The values *conformity* and *tradition* were reflected in the nursing staff's conservativeness and preference to maintain their traditional routines. The nursing staff mentioned that altering their care routines was more difficult than expected. They tended to continue doing their rounds and checking on residents as usual despite the use of new technologies [49]. Managers' proxy perspectives also recognized the difficulties in altering care routines [48]. Although nursing staff tended to keep their traditions, they were generally supportive of new technologies that contributed to improvements in daily practice, particularly when the technologies functioned as intended [46]. However, new surveillance technologies never functioned properly from the outset, and implementing them often resulted in initial malfunctions [50]. These initial malfunctions, such as false alarms, poor Wi-Fi, and slow software had a negative influence on the level of acceptance, partly due to the nursing staff's great sense of responsibility for the residents [7,46,51]. These challenges limited the usability and accuracy of detecting unsafe situations and led to an even higher workload [7,22,46,48]. In addition, these challenges resulted in nurses' alarm fatigue and, as a result, nursing staff primarily relied on their previous experiences with residents' routines [21,48,51].

Moreover, nursing staff found new technologies challenging because these technologies required skills they did not previously need in their daily practice [52]. The extent to which procedures and instructions were tailored to the nursing staff's daily work, their shifts, and professional language influenced the nursing staff's openness to changes and the experienced extent of *self-direction* [7,20,51,53]. Moreover, they expected the technology vendors to help them, for instance, by providing support even outside regular business hours [7,51]. Nursing teams also felt supported when some of their colleagues took the lead in teaching them [7,47,51].

Although surveillance technologies contributed to providing optimal security and safety for residents, nursing staff expressed concerns about the impact of the use of monitoring technology on privacy, their competence in using the technologies, and the replacement of their roles [51,54]. This outcome relates to the value *self-direction.* Nursing staff expressed concerns that the monitoring technology was applied as a "big brother tool," indicating a lack of confidence [7,19,37,54]. Nursing staff preferred the application of codes of ethics and limited access by authorized professionals to protect their privacy and that of residents [7,53]. In addition, staff were concerned that monitoring had an impact on resident relationships [54] and that their role would be replaced by technology [51]. The latter was also recognized by care managers, as reflected in their proxy perspectives [51,52]. Particularly, older nursing staff expressed concerns about experiencing challenges while working with technology, resulting in a reserved attitude toward care technology [51]. The extent to which nursing staff felt involved influenced the freedom they experienced to determine their ideas and actions. This outcome impacted their openness to changes, represented by the values *self-direction in action and thought*. From a proxy perspective, care managers and administrators recognized these challenges faced by nursing staff, and they mentioned that nursing staff needed time to adapt to new care technologies before they could appreciate them [55,56].

Because surveillance technologies directly interfered with the nursing process, nursing staff expressed a desire to be involved from the beginning, for instance by being involved in discussions and decision-making processes [22], to be able to express their needs, opinions, and concerns [53]. The values *achievement* and *power* are related to the nursing staff's desire to be acknowledged and appreciated as important stakeholders.

In summary, the values *benevolence, security, conformity,tradition*, *self-direction in action and thought*, *achievement,* and *power* were represented among the stakeholder group of nursing staff.

Residents

Of the 31 included articles, 20 (65%) described residents as stakeholders, with 17 (85%) articles based on a (partial) proxy perspective of nursing staff (n=12, 47%), informal caregivers (n=4, 20%), LTC administrators (n=2, 10%), and developers of surveillance technologies (n=2, 10%).

Several residents who were interviewed, said they were aware of their dependency on care as a result of their cognitive decline. They adapted to life's circumstances, reflecting the value *humility*. Moreover, they indicated that care customized to their preferences and needs supported their independence and

contributed to their safety [52,55]. The value *(personal) security* is represented in the residents' feeling of being heard by care professionals. This means that professionals know where they are, respond to their alarms, and can care for them [22,53,57]. Consequently, surveillance technologies were perceived as part of the deal, improving their safety, and receiving individualized care [38,57]. This was recognized by nursing staff because they mentioned that surveillance technologies enhanced safety and contributed to the care, in line with residents' personal preferences, contributing to the latter's level of independence [52,53,55].

Residents knew that they partly gave up their privacy when they moved to a nursing home. Nevertheless, surveillance technologies could contribute to their feeling of *(personal) security* because the technologies protect their privacy. Coded doors, for example, could prevent other residents from unintentionally entering their rooms [19,38]. Although residents expressed a feeling of increased *(personal) security* due to surveillance technologies, they feared the consequences of these technologies as they expressed worries that these technologies would replace the valuable human contact with staff, because social contacts were an essential and highly valued part of life [37,38].

Furthermore, the values *conformity* and *tradition* seemed important to residents. These values include the avoidance of significant changes in their living environment, for instance, due to the use of surveillance technologies. Residents emphasized the importance of maintaining a feeling of homeliness [38,57]. In addition, surveillance technology should not jeopardize their feeling of homeliness or be (too) visible. Instead, surveillance technology should be aesthetically pleasing, easy to use, and not disrupt their daily routines [57,58]. Caregivers, as proxies, noticed that the extent of devices' intrusiveness to residents influenced the acceptability of devices [47,55,59]. Moreover, several residents expressed resistance to technology based on the usefulness they experienced [60]. The relevance of these factors was recognized in the proxy perspective of nursing staff and informal caregivers [47,59,61]. In addition, nursing staff experienced fewer nighttime disturbances for residents resulting in calmer nights [7,51]. Nursing staff noticed that the residents' openness to changes and the experienced *self-direction* decreased when there were more false alarms [57], which was recognized in the reluctance the latter expressed and is related to the experienced usefulness.

Residents' feelings of being stigmatized or being regarded as patients increased when their wishes regarding the visibility, appearance, and usability of surveillance technologies were not met [18]. This outcome aligns with the value *face*, which emphasizes maintaining one's public image and avoiding humiliation. Regarding this value, residents expressed greater concern about cameras than about other devices. These concerns were particularly about being recognizable in images while performing personal and hygienic activities, evoking feelings of intrusion and vulnerability [37,38,54]. This factor was similarly mentioned by managers when they expressed their concerns about the invasion of residents' privacy and dignity arising from surveillance cameras [54].

In summary, the values *humility, (personal) security, conformity, tradition, self-direction, and face* were most represented in the resident stakeholder group.

Informal Caregivers

Of the 31 articles, 13 (42%) mentioned informal caregivers (such as family caregivers and authorized representatives) as stakeholders. In total, 7 (53%) of these articles (partly) described a proxy perspective. Proxy perspectives were mostly represented by nursing staff (n=6, 46%), LTC administrators (n=2, 15%), and physicians (n=1, 8%).

Informal caregivers, mostly family members, expressed their concerns about the safety of their loved ones. Preventing (new) falls was often mentioned as a reason to use surveillance technology [19,22]. In their desire to contribute to the safety and well-being of their loved ones, deriving from the values *benevolence* and *security*, informal caregivers often valued residents' personal safety above possible threats to their privacy and freedom of movement [22,38]. This was also recognized by nursing staff and physicians, who observed informal caregivers' peace of mind when surveillance technologies were used [21,54]. In line with this value benevolence, informal caregivers felt responsible for the well-being of their loved ones. Informal caregivers feared that surveillance technologies would replace valuable human contact that arises from this value, and this fear may be reinforced by informal caregivers' awareness of the staff shortages in nursing homes [38,51]. This fear was recognized by nursing staff who indicated that they perceived it among informal caregivers [38,51]. Informal caregivers noted surveillance technologies should support nursing staff rather than replace them [38]. From a transcendent perspective, the value *universalism* seemed to be represented in informal caregivers' concerns about striving for equality and protection of people who were weak [31].

Informal caregivers mentioned they were willing to accept a wide range of surveillance technologies, including video surveillance, as long as they were convinced about the contribution these technologies made to the safety, quality of life, and well-being of their loved ones [22,37,38]. In addition, nursing staff and managers said that the level of informal caregivers' openness to changes was also determined by their perception of usefulness and their interest in technologies [53]. Informal caregivers' willingness to accept a broad range of surveillance technologies and their openness to changes reflected the value *self-direction*, embodying their ability to choose their goals and be involved in decision-making. The prerequisite of getting involved in (discussions about) applying surveillance technologies to their loved one [38,47,53,62] arises from the value *power*. Informal caregivers mentioned that these discussions should occur between all relevant stakeholders, such as residents, relatives, and nursing staff. In addition, informal caregivers wanted to be asked for formal consent as they were (authorized) representatives [38]. In practice, informal caregivers mentioned that they were not or not sufficiently informed about the available surveillance technologies [62].

In summary, the values *benevolence* and *security, universalism, self-direction, and power* were represented in the stakeholder group of informal caregivers.

Care Managers

Of the 31 articles, 9 (29%) mentioned care managers as stakeholders. Several care managers mentioned that their priority was to manage the 24×7 health care service and that surveillance technologies could contribute to achieving this aim [53,54]. The application of surveillance technologies increased the safety of residents, which is related to the values *benevolence* and *security*. Furthermore, data from these systems could be used to defend nursing homes against allegations of negligence leveled by families [19]. In line with this situation, care managers mentioned that data could help them monitor staff and hold them to account. However, care managers acknowledged that responding to an alarm was no guarantee that care was being provided [19]. The ethical objections against surveillance technologies that care managers mentioned were particularly aimed at the potential impact on residents' privacy, rather than the impact on nursing staff [19,22]. They acknowledged camera surveillance could contribute to a "big brother" effect and a culture of mistrust [19,54]. For care managers, surveillance technologies particularly seemed to represent values with a personal focus, namely, to have control over and manage the residential care facility. This outcome aligned with values such as *power* and *achievement.*

Care managers mentioned they were often insufficiently prepared for new ways of working and, subsequently, the different authority structures resulting from implementation strategies. In addition, they mentioned they were unable to make implementations a priority due to other organizational priorities [51]. The unpreparedness for changes they experienced could be a consequence of the changed ways of working, fear of the unknown, and clinging to the values *conformity and tradition*. Nevertheless, they were open to changes, although they also experienced unpreparedness for cocreation from several stakeholders [51]. This unpreparedness stems from an expectation of a tailored solution from vendors—not a realization that everyone's input including their own was a prerequisite for a joint implementation process [51]. Care managers taking the initiative in prioritizing reflections with other stakeholders to discuss dilemmas was mentioned as a facilitator for the successful use of surveillance technologies [51]. In this respect, care managers faced challenges in their *self-direction in action and thought* [31].

Furthermore, care managers felt that values *power* and *achievement* might be occasionally threatened. They indicated that they were concerned about increasing costs associated with new technologies, such as surveillance systems, while revenues remained stagnated [63]. In addition, surveillance technologies were not always as robust as they needed to be to withstand use in nursing home practice, leading to recurring costs due to damaged products [19]. Furthermore, care managers mentioned they felt restricted by rules and contractual obligations with vendors, which could hinder their access to technologies [19].

In summary, the values *benevolence* and *security* were somewhat represented in the stakeholder group of care managers. However, the values *power,achievement, conformity,tradition,* and *self-direction in action and thought* were more clearly represented.

Developers

A total of 7 (23%) of the 31 articles mentioned developers as stakeholders. Developers envisioned surveillance technologies would be used to enhance the safety of residents with dementia and improve the security of residents and nursing staff while respecting the privacy of both [19,55,58]. This outcome aligned with the value *security*.

In addition, developers mentioned they sought to use their technologies to enable nursing staff to support and assist residents with progressive diseases such as dementia. Due to the characteristics of the resident population, developers mentioned that surveillance technologies should be dynamic and scalable and be designed for failure and intensive use [55]. This outcome reflected the value *benevolence* as it underscored their determination to support and assist those in need.

Developers emphasized that by testing monitoring technologies new insights were created and improvements could be made. Testing in real life supported them to achieve success, in line with the values *achievement* and *power*. However, they acknowledged that high error rates in initial tests had an influence on the nursing staff's openness to change [52,58,60,61]. Conversely, many initial errors occurred due to unskillful use. A higher level of training for nursing staff could reduce these errors [58]. This outcome challenged developers' values *stimulation* and *self-direction* because they were generally excited about new technologies when they noticed the impact of their technologies on end users in nursing homes [60]. Furthermore, developers emphasized the importance of evaluating the effects of surveillance technologies in nursing homes [60]. Moreover, they underlined the importance of development in close collaboration with end users and specialists in dementia care to meet their needs and requirements [52,55]. Developers acknowledged that this close collaboration was an intensive process [52,55].

In summary, for the developer stakeholder group, the values *security, benevolence, achievement, power, stimulation,* and *self-direction* were most represented.

Physicians

Of the 31 articles, 5 (16%) included physicians as stakeholders in the application of surveillance technologies.

As with nursing staff and informal caregivers, the values *security* and *benevolence* were represented, as physicians agreed that providing safety was an important reason to apply these technologies. Moreover, they noted these technologies offered peace of mind to nurses and informal caregivers [21,37]. This outcome might explain physicians' high acceptance rate of surveillance technologies [55,58].

In summary, the values *security* and *benevolence* were clearly represented in the stakeholder group of physicians.

LTC Administrators

LTC administrators were mentioned in 4 (13%) of the 31 articles. LTC administrators expressed their vision of being at the forefront of implementing new technologies, especially in newly constructed nursing homes. Generally, new technologies

were considered an important solution for health care challenges [52]. In these administrators' perspectives, the values *self-direction in action and thought* and *stimulation* were reflected. However, LTC administrators acknowledged that implementing these technologies challenged institution's openness to changes because these new technologies demanded time and resources across various roles and professions [51]. LTC administrators noted that education for staff was an important issue; however, they sometimes questioned the abilities of nursing staff to master the new technologies despite offering education to the latter [37].

The values *power* and *achievement* emerged in the acknowledgment by LTC administrators that their residential care facilities were perceived as more attractive employers when they used modern surveillance technologies [37,52]. The financial investments required could force organizations to opt for a cheaper but more generic design, although this may be less suitable than preferred designs [19]. Consequently, LTC administrators could feel restricted in their control over resources. In line with care managers' perspective, LTC administrators mentioned that organizational contracts limited their scope and flexibility around product choices and ongoing maintenance [19].

In summary, the values *self-direction in action and thought, stimulation, power,* and *achievement* were represented in the stakeholder group of LTC administrators.

Values of the Other Stakeholders

The other stakeholder groups, that is, maintenance staff, ICT staff, vendors of surveillance technologies, project managers, and academics were mentioned only once or twice in the 31 articles included. Maintenance staff mentioned that although new surveillance technologies might seem to have a limited scope, only affecting night shift workers and residents, these technologies could also have an influence on janitors, cleaning staff, and substitute personnel [37,51]. In their view, maintenance staff members were also in need of information and education to accommodate the technologies in their (cleaning and maintenance) routines. When they were not informed or did not receive education, they reported difficulties, for example, in replacing sensors and reconnecting cables after cleaning [37,51]. Because a new system is only as good as the people who are responsible for operating it, the people operating the system could affect its reliability [37,51]. Maintenance staff expressed a desire to be acknowledged as important stakeholders and be involved in education, which relates to their openness to changes in the values *self-direction* and *stimulation*. In line with the apprehensions of nursing staff, maintenance staff expressed concerns about being observed [37]. For instance, maintenance staff sometimes felt threatened due to their openness to changes and *self-direction*, reducing their motivation to engage with new technologies.

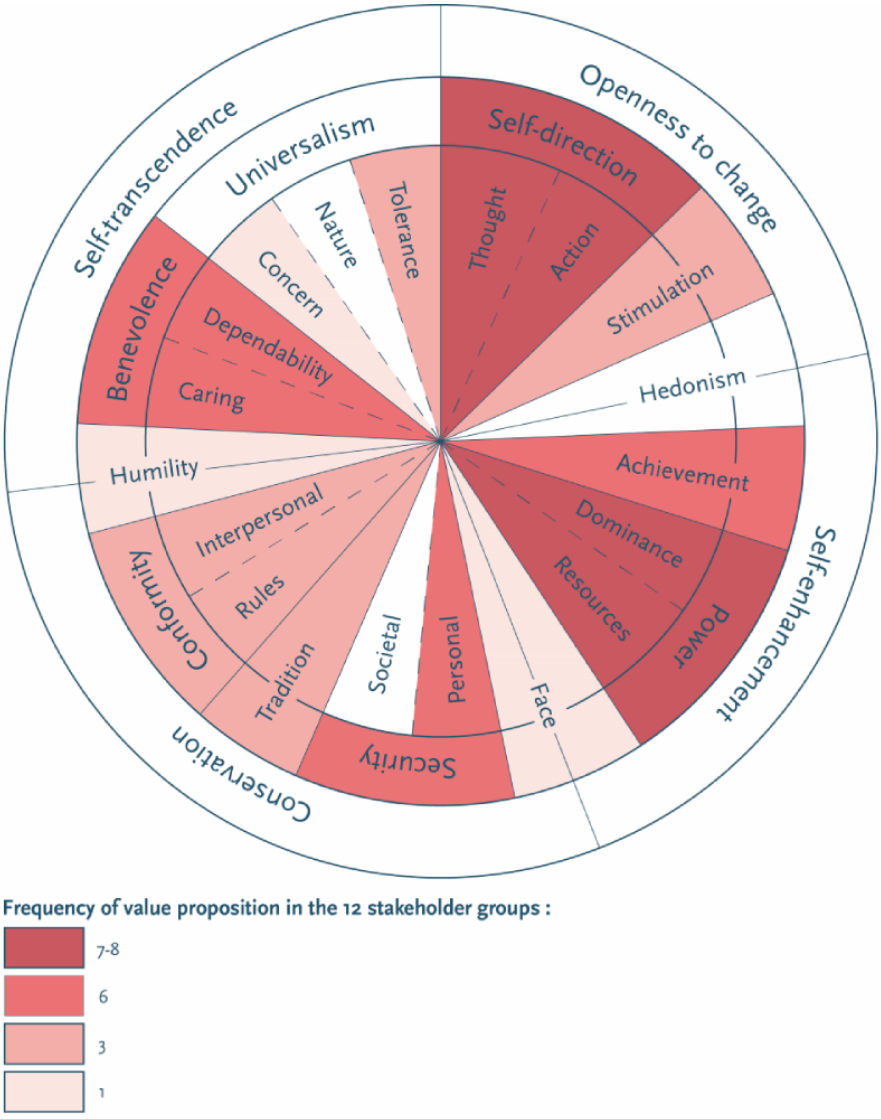
ICT staff emphasized the importance of system and component interoperability for ensuring system reliability because new systems were often installed into existing systems and infrastructure [19,51]. However, they mentioned that this interoperability between systems often was not facilitated by manufacturers [19]. Furthermore, the staff highlighted the importance of their involvement from the outset in exploring this compatibility and interoperability to prevent compromises in residents' safety and security [51]. This outcome reflected the need for ICT staff to be involved and acknowledged as stakeholders, in line with the values *power* and *achievement*. However, in practice, it was seen that ICT staff only became seriously involved when systems were unstable or errors occurred [51].

In the research by Dugstad et al [51], vendors of surveillance technologies and ICT were challenged to adopt a more socially focused approach to bridge differences between stakeholder groups. The vendors mentioned that despite their knowledge about their products, they needed the nursing staff's insights to ensure their technologies worked in specific care environments due to the great variety of care practices and infrastructure. Besides vendors striving for their own success, reflecting the values *power* and *achievement*, they were challenged to pursue a higher purpose, in line with the value *universalism*. Universalism relates to vendors' willingness to engage with other stakeholders to contribute to successful applications. In line with this value, ICT staff and vendors recognized being challenged to adopt a language that was more understandable for nursing staff who experienced differences in professional cultures and language (jargon) [51]. The ICT staff and vendors mentioned they were not used to adapting their language; therefore, misunderstandings were initially manifested due to a lack of knowledge and insight into each other's workflows. This situation challenged them to transcend their profession and jargon in order to contribute to a successful digital transformation [51]. Local project managers also underlined the importance of recognizing and bridging cultural differences as conditions for ensuring successful digital transformation. Hence, project managers identified their role as translators between different stakeholders [51], in line with the value *universalism*.

Finally, academics mentioned that the balance between freedom and security is important. They tended to value freedom of movement over security [20]. The socially focused value *benevolence* was reflected in being devoted to the welfare of others and protecting people with an increased care dependency. However, academics tend to attach great importance to residents' freedom to determine their own actions despite the dementia, which relates to the *value self-direction*.

Figure 2 presents the stakeholders' basic human values. The darker the color scheme, the more often this value is applied to various stakeholders.

Figure 2. Value palette for all stakeholders in the proposed circular structure of the 19 basic human values by Schwartz [32].



Discussion

Principal Findings

This scoping review aimed to identify the stakeholders involved in applying surveillance technologies for people with dementia in nursing homes and to describe the values of these stakeholders. Overall, 12 stakeholder groups were identified in 31 articles. The stakeholder groups of nursing staff, residents, and informal caregivers were most often mentioned in these articles. Several stakeholder groups, such as physicians, LTC administrators, maintenance staff, ICT staff, vendors, and project managers, were mentioned significantly less frequently [51,52,55].

Although many stakeholder groups emphasized the importance of being acknowledged and feeling involved as stakeholders [22,37,51,53,62], they said that their involvement should be improved [22,51,53,62]. Moreover, several stakeholders mentioned feeling dissatisfied when they did not feel sufficiently involved [53,62]. As dissatisfaction could detract from success, all relevant stakeholders should be involved [53]. Furthermore, in literature, the importance of involving both primary stakeholders, also known as end users, and secondary stakeholders (ie, more distantly involved stakeholders) is underlined as contributing to the successful implementation of care technologies, such as surveillance technologies [22,28,29].

To determine which primary and secondary stakeholders should be involved and engaged, it is important to identify these stakeholders [69]. Identifying stakeholders can be accomplished, for example, through a stakeholder analysis [69]. Several stakeholder matrices such as the power-interest matrix, the 3D matrix, and the responsible, accountable, consulted, and informed matrix offer tools to identify stakeholders and categorize their attributes such as their power, position, and level of interest [69,70]. In addition, these matrices support prioritizing who should be involved and to what extent. These matrices reflect the increased recognition of how the characteristics of stakeholders influence innovation and implementation processes [69,71,72]. Regardless of the distinction between stakeholders' level of involvement and their responsibilities, the resistance of any stakeholder can have an influence on the success of the implementation of care technologies.

Resistance by stakeholders is a challenge, necessitating the identification of alternative perspectives to a situation [73-75]. Therefore, besides identifying stakeholders when applying care technology, it is necessary to realize that care technologies strongly influence the care process and everyone involved. Therefore, cooperation between stakeholders who have previously rarely cooperated is essential. Hence, it is important to thoroughly know and acknowledge the different stakeholders and their differences and respect their differing interests [29].

Several implementation theories provide theoretical support to implementation processes, especially in complex contexts such as health care. One of the theoretical constructs for involving multiple stakeholders in implementation processes in health care is the normalization process theory (NPT) by May et al [76]. NPT provides tools to enhance an understanding of the social processes of thinking, enacting, and organizing work to implement and adopt interventions in care processes within health care organizations [77]. NPT distinguishes 4 constructs: coherence, cognitive participation, collective action, and reflexive monitoring. These 4 constructs emphasize the relevance of knowledge about the value that stakeholders assign to care technology and attention to stakeholders' willingness to participate and cooperate. NPT also focuses on keeping people engaged during the whole implementation process, including reflecting upon and appraising the effect of newly implemented technology [29,78].

A common thread across the NPT as well as other implementation theories is the need for cooperation and communication between stakeholders focusing on knowing and acknowledging each other's perspectives, values, and interests [29,79]. For instance, in the early 90s, Gregory and Keeney [72] wrote about the right of multiple stakeholders to be involved in policy decision processes and consequently balance between conflicting objectives. In addition, they mentioned that values, beliefs, cooperative potential, and stakeholders' concerns are known to influence the outcome of innovation processes [72]. Similarly, in business ethics, Dentoni et al [80] mentioned the complexity of multiple stakeholder involvement in dynamic settings because each stakeholder group has its own set of values, perceptions, and interests that may clash. This situation requires fundamentally different approaches that demand understanding differing values, complex settings, and acting upon uncertain knowledge [80].

The active participation of stakeholders challenges them to effectively collaborate with a critical and open stance toward their perspectives and values. Several stakeholders in this scoping review mentioned that this collaboration demands not only time and effort but also challenges them to adapt their jargon to interprofessional cultures and bridge their differences [51,52]. Although all stakeholders should be able to operate and communicate across boundaries between different practices with each other, collaborating with people is difficult and can lead to tensions and misunderstandings related to values and interests [81]. Such collaboration demands competence to perceive differences as learning opportunities and to cross boundaries between multiple stakeholders [81,82]. Collaboration between stakeholders is receiving increasing attention from organizations. In addition, organizations have shown a growing interest in creating value through participation and interaction with multiple stakeholders. However, until now, most of the attention has been given to creating value *for* stakeholders and not *with* them. In cases of increased awareness about stakeholders' input, values and interests may be well identified [83]. Thus, besides merely identifying and superficially involving stakeholders, it is important to pay attention to their perspectives, values, and interests.

Therefore, this scoping review also focused on what is already known about the stakeholder group values identified in the 31 articles reviewed. Several values in the theory of basic human values by Schwartz et al [31] were frequently represented among the stakeholder groups. The values *benevolence* and *security* were represented in 6 (50%) out of 12 stakeholder groups. This outcome is unsurprising given the progressive nature of dementia. Most stakeholder groups mentioned they experienced a feeling of being responsible for caring for residents with dementia and responding to their needs, in line with the value *benevolence* [21,22,37,38,55]. Concerning the value *security*, striving for safety for residents with dementia is often mentioned as a reason to apply surveillance technologies [7,22,46]. However, the meaning assigned to this value varies among the different stakeholder groups. The trade-off between safety and aspects such as privacy and freedom of movement differs among various stakeholders and appears to be related to how closely a stakeholder is involved with a resident [20]. For nursing staff, *security* is related to enhancing the safety of residents and mitigating their risk of falling [7,22,46]. This relates to the responsibility they feel for their residents because they feel accountable for accidents or injuries of residents [22,48]. For care managers, the value *security* is also related to their accountability for providing care, managing the residential care facility, assuring families that care is provided, and monitoring staff [19,54]. For residents, *security* is related to the feeling of being heard regarding their personal preferences and (care) needs and experiencing a feeling of homeliness [22,52,53,57]. Informal caregivers have concerns about their loved ones and they seem to experience more peace of mind when surveillance technologies are applied [21,54]. Consequently, informal caregivers often value safety above freedom of movement and the possible threats to privacy [19,22]. In contrast, academics, a group of stakeholders who do not have a close relationship with residents, tend to value freedom of movement above safety, giving high importance to the residents' experience of freedom and self-determination [20].

As is evident in the considerations regarding safety, surveillance technologies that are applied in practice affect the work and living environment of several stakeholder groups [37,51,53]. Accordingly, many stakeholders emphasized their desire to be and to feel involved in the application of surveillance technologies [22,53]. Arising from this desire, the values *self-direction in action and thought* were represented in 8 (67%) of 12 stakeholder groups. This outcome underlines the relevance of thoroughly involving stakeholders throughout the implementation and application process. In addition, the degree of stakeholders' openness to changes, their experienced *self-direction*, and their tendency to cling to traditional routines seem to be related to the extent to which they feel involved.

The more they feel involved, the more they are open to changes and willing to collaborate [22,38,47,53,62]. Being and feeling involved is especially an important issue for nursing staff because surveillance technologies directly interfere with the nursing process. Therefore, nursing staff were the most frequently cited stakeholders. They were mentioned in 23 (74%) of the 31 articles. Accordingly, they expected to be involved from the beginning; asked to express their needs, opinions, and concerns [53]; involved in decision-making; and acknowledged as important stakeholders [22]. Informal caregivers mentioned that they were willing to accept a broad range of technologies for their relatives as long as the former were informed about the surveillance technologies, were involved in deliberations about the technologies, and were asked for their consent regarding the application of the technologies [38,47,53,62].

Related to the value *self-direction* were the values *power* and *achievement*, representing stakeholders' desire to exert their influence on other people or use material resources and pursue success in competencies and performance. The values *power* and *achievement* were mentioned by 7 and 6 stakeholder groups, respectively. These values are a reflection that stakeholders such as nursing staff, informal caregivers, developers, ICT staff, and vendors want to feel recognized in their knowledge and experiences, and want to exert their influence in the application of surveillance technologies [22,38,51-53,61,62]. For care managers and LTC administrators, the values *power* and *achievement* are related to control over and the management of residential care facilities [53,54].

Although several values such as benevolence, security, self-direction, power, and achievement were represented by most of the stakeholders, the various stakeholder groups assigned different meanings to these values. Knowing and understanding diverse stakeholders' perspectives and attitudes, including the different meanings they attribute to values, is crucial because this knowledge and understanding influence the adoption and use of technologies [84]. Integrating multiple perspectives is valuable to fully understand the complexities of care practices [29] and dementia care technology [84]. In addition, it is challenging to distinguish whether stakeholders give meaning to a certain value based on their interests or whether they act from the resident's perspectives. This situation demands that designated persons in health care organizations have courage and take the lead in initiating meaningful and in-depth conversations where the diverse stakeholders will be challenged to communicate across their boundaries, looking beyond their own perspectives.

Methodological Considerations

A proxy perspective was often observed in the stakeholder groups of nursing staff, residents, and informal caregivers. This outcome is not unusual because proxy perspectives often originate from stakeholders with whom one works or lives [85]. In addition, proxy perspectives are often seen in stakeholder groups that collaborate and deal with matters that touch on values and interests. Hence, a proxy perspective was regularly seen among nursing staff and informal caregivers in this scoping review. Furthermore, when making assumptions, people are less understanding of others' actual motivations [85]. Accordingly, Kloos et al [86] researched the well-being of residents and found that nursing staff tended to overestimate the well-being of residents. Their study underlines the importance of combining proxy assessments with self-reports whenever possible. In addition, Kunicki et al [87] noted that the level of involvement of proxies and their sense of well-being could influence their perception of the resident's preferences. Kunicki et al [87] recommended that proxies find methods to better understand residents' preferences when residents were not able to express their preferences properly [87]. This scoping review raises the question of whether proxy perspectives of residents with dementia reveal resident's perspectives. Although it is challenging to interview residents with dementia due to the impairments related to their condition, it is possible to involve them. Therefore, in future research, residents should be made participants in research to understand their perspectives [88,89].

Strengths and Limitations

Previous studies primarily focused on facilitating and limiting factors and ethical dilemmas when applying surveillance technologies to people with dementia [18,23,36-38]. However, this scoping review is the first to explore which universally recognized human values are reflected in the experiences and opinions of stakeholders. Because values reflect human thinking and determine attitude and behavior, it is crucial to consider the underlying values that explain stakeholders' behavior and reactions [34,35,79]. Stakeholders' values were classified using the basic human values model by Schwartz. Although this model offers an empirically tested theory to predict attitudes and behaviors in different contexts and situations, it is possible that not all opinions and experiences of stakeholders could be categorized using the model. Furthermore, it is likely that the stakeholders' values do not provide a complete picture of the values that stakeholders hold in practice because the representation of stakeholders' values in this scoping review is based on information that did not primarily focus on stakeholders' values.

Recommendations for Research

Although this scoping review identified 12 stakeholder groups, most (26/31, 84%) of the articles included in the review described only 4 or fewer stakeholder groups. Therefore, various stakeholder groups were underrepresented. Future research should emphasize the involvement of all relevant stakeholders. In addition, our explorations of stakeholders' values revealed there is insufficient information about stakeholders' values. Hence, more research about the values that influence stakeholders' actions and decisions should be conducted. In exploring stakeholders' perspectives and values, a proxy perspective should be avoided where possible. In future research, inviting stakeholders to look beyond their perspectives and boundaries could be useful in mitigating language and knowledge boundaries between different stakeholders. This approach could facilitate constructive cooperation between stakeholders. In addition, efforts should be made to include residents with dementia in research to explore their perspective rather than assumptions about residents' perspectives being largely based on a proxy perspective. Listening to people with dementia can enhance the quality, relevance, and impact of

dementia research, which contributes to the enhancement of knowledge based on what we learn *from* them and their informal caregivers, in order to create knowledge *with* them [90]. Participatory research could apply to this kind of research about complex dynamic subjects; research should be conducted with rather than about stakeholders.

Conclusions

All stakeholders involved in applying surveillance technologies expressed a desire for their perspectives and values to be acknowledged. This desire stems from the human need to be acknowledged and appreciated. Moreover, all the stakeholders manifested a willingness to be engaged and participate. The broad acknowledgment and involvement of stakeholders and an understanding of their perspectives and values contribute to the successful implementation and application of surveillance technologies for people with dementia in nursing homes. Therefore, when applying surveillance technologies for people with dementia, residential care facilities are expected to intensively collaborate with an increasing number of stakeholders. Therefore, stakeholders' active engagement, with attention to everyone's perspectives and values, is more important than ever.

Acknowledgments

The authors thank information specialist C Niehot for the support in drawing up the search string for this scoping review. In addition, the authors thank Professor Dr M Verkerk for the conversations about values, interests, and ideals and Professor Dr S Schwartz, Drs D Arets, R van de Vorst, and Dr B Wernaart for the helpful feedback in using the theory of values by Schwartz et al [31] to organize the results of this scoping review according to human values.

Authors' Contributions

All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 167 KB - aging_v8i1e64074_app1.pdf]

Multimedia Appendix 2
Overview of the search string for the 5 databases.
[DOCX File , 22 KB - aging_v8i1e64074_app2.docx]

Multimedia Appendix 3
Mapping overview of the included articles.
[DOCX File , 73 KB - aging_v8i1e64074_app3.docx]

References

1. Ageing and health. World Health Organization. 2024 Oct 1. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-12-12]
2. Dementia in Europe Yearbook 2019: estimating the prevalence of dementia in Europe. Alzheimer Europe. 2019. URL: https://www.alzheimer-europe.org/sites/default/files/alzheimer_europe_dementia_in_europe_yearbook_2019.pdf [accessed 2023-05-04]
3. Global action plan on the public health response to dementia 2017 - 2025. World Health Organization. 2017 Dec 07. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2023-05-04]
4. Huijsman R, Boomstra R, Veerbeek M, Dopp C. Zorgstandaard dementie. Dementiezorg voor Elkaar. 2020. URL: https://www.zorgstandaarddementie.nl/zorgstandaard [accessed 2021-09-21]
5. Strengthening the dementia care workforce: a public health priority. American Public Health Association. 2020 Oct 24. URL: https://www.apha.org/policies-and-advocacy/public-health-policy-statements/policy-database/2021/01/13/strengthening-the-dementia-care-workforce [accessed 2023-12-18]
6. Global strategy on human resources for health: workforce 2030. World Health Organization. 2020 Jul 07. URL: https://www.who.int/publications/i/item/9789241511131 [accessed 2023-05-04]
7. Emilsson M, Karlsson C, Svensson A. Experiences of using surveillance cameras as a monitoring solution at nursing homes: the eldercare personnel's perspectives. BMC Health Services Research 2023 Feb 10;23(1):144 [FREE Full text] [doi: 10.1186/s12913-023-09130-2] [Medline: 36765321]

8. Zorg voor de toekomst: over de toekomstbestendigheid van de zorg. Sociaal-Economische Raad. 2020 Jun. URL: https://www.ser.nl/-/media/ser/downloads/adviezen/2020/zorg-voor-de-toekomst.pdf [accessed 2023-05-04]
9. Bijleveld C, Boot A, Bovens M, Engbersen G, Hulscher S. Kiezen voor houdbare zorg. Wetenschappelijke Raad voor het Regeringsbeleid. 2021. URL: https://www.wrr.nl/publicaties/rapporten/2021/09/15/kiezen-voor-houdbare-zorg [accessed 2022-05-01]
10. Dementia statistics hub. Alzheimer's Research UK. URL: https://dementiastatistics.org/ [accessed 2023-04-05]
11. Mjørud M, Selbæk G, Bjertness E, Edwin TH, Engedal K, Knapskog AB, et al. Time from dementia diagnosis to nursing-home admission and death among persons with dementia: a multistate survival analysis. PLoS One 2020 Dec 4;15(12):e0243513 [FREE Full text] [doi: 10.1371/journal.pone.0243513] [Medline: 33275638]
12. Factsheet dementie. Vektis. 2022 Jan 24. URL: https://tinyurl.com/59xxymzn [accessed 2022-01-27]
13. Bauer AK, Fassmer AM, Zuidema SU, Janus SI, Hoffmann F. End-of-life care in German and Dutch nursing homes: a cross-sectional study on nursing home staff's perspective in 2022. Arch Public Health 2024 Jun 14;82(1):85 [FREE Full text] [doi: 10.1186/s13690-024-01316-2] [Medline: 38877598]
14. Nap HH, van der Weegen S, Cornelisse L, Lukkien D, van der Leeuw J. Zorgdomotica in de nachtzorg: onderzoek naar de mogelijkheden en onmogelijkheden van de nieuwe generatie zorgdomotica. Vilans. 2015. URL: https://www.researchgate.net/publication/287216465_Zorgdomotica_in_de_nachtzorg [accessed 2023-05-04]
15. Stroomversnelling voor krappe arbeidsmarkt in verpleeghuiszorg – uitgebreide rapportage. Ministerie van Volksgezondheid, Welzijn en Sport. 2020 Dec. URL: https://www.igj.nl/publicaties/rapporten/2020/12/10/stroomversnelling-voor-krappe-arbeidsmarkt-in-verpleeghuiszorg-%E2%80%93-uitgebreide-rapportage [accessed 2023-05-04]
16. van der Roest H, van der Kuil M, Overbeek A, Hartstra E. 414 - Supporting staff delivering person-centered care to people with dementia in Dutch nursing homes. Int Psychogeriatr 2021 Oct;33(Supplement 1):37-38. [doi: 10.1017/s1041610221001733]
17. Mohammadnejad F, Freeman S, Klassen-Ross T, Hemingway D, Banner D. Impacts of technology use on the workload of registered nurses: a scoping review. J Rehabil Assist Technol Eng 2023 Jun 17;10:20556683231180189 [FREE Full text] [doi: 10.1177/20556683231180189] [Medline: 37342268]
18. Niemeijer AR, Depla MF, Frederiks BJ, Hertogh CM. The experiences of people with dementia and intellectual disabilities with surveillance technologies in residential care. Nurs Ethics 2015 May;22(3):307-320. [doi: 10.1177/0969733014533237] [Medline: 24913545]
19. Hall A, Brown Wilson C, Stanmore E, Todd C. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. BMC Geriatr 2019 May 24;19(1):145 [FREE Full text] [doi: 10.1186/s12877-019-1155-6] [Medline: 31126240]
20. Niemeijer AR, Frederiks BJ, Depla MF, Legemaate J, Eefsting JA, Hertogh CM. The ideal application of surveillance technology in residential care for people with dementia. J Med Ethics 2011 May;37(5):303-310. [doi: 10.1136/jme.2010.040774] [Medline: 21292695]
21. Zwijsen SA, Depla MF, Niemeijer AR, Francke AL, Hertogh CM. Surveillance technology: an alternative to physical restraints? A qualitative study among professionals working in nursing homes for people with dementia. Int J Nurs Stud 2012 Feb;49(2):212-219. [doi: 10.1016/j.ijnurstu.2011.09.002] [Medline: 21962335]
22. Hall A, Wilson CB, Stanmore E, Todd C. Implementing monitoring technologies in care homes for people with dementia: a qualitative exploration using Normalization Process Theory. Int J Nurs Stud 2017 Jul;72:60-70 [FREE Full text] [doi: 10.1016/j.ijnurstu.2017.04.008] [Medline: 28494333]
23. Niemeijer A, Depla M, Frederiks B, Hertogh C. Toezichthoudende domotica: een handreiking voor zorginstellingen. VUmc. 2012. URL: https://nvavg.nl/wp-content/uploads/2014/03/2012-Toezichthoudende-domotica-een-handreiking-voor-zorginstellingen.pdf [accessed 2022-04-12]
24. Wet zorg en dwang. Ministerie van Volksgezondheid, Welzijn en Sport. 2019. URL: https://www.dwangindezorg.nl/wzd [accessed 2021-05-28]
25. Handreiking Wet zorg en dwang voor specialisten ouderengeneeskunde & artsen verstandelijk gehandicapten in intramurale setting. Verenso. 2022 Jun. URL: https://www.verenso.nl/_asset/_public/Richtlijnen_kwaliteit/richtlijnen/Handreiking-Wzd-SO-en-AVG_versie-2-0-_DEFINITIEF.pdf [accessed 2025-02-18]
26. Recommendations on how to improve the legal rights and protection of adults with incapacity due to dementia. Alzheimer Europe. 2017. URL: https://tinyurl.com/3w9uby7f [accessed 2023-04-05]
27. Greenhalgh T. How to improve success of technology projects in health and social care. Public Health Res Pract 2018 Sep 27;28(3):2831815 [FREE Full text] [doi: 10.17061/phrp2831815] [Medline: 30406256]
28. Wernaart B. In: Wernaart B, editor. Moral Design and Technology. Wageningen, The Netherlands: Wageningen Academic Publishers; 2022.
29. Wouters EJ. Technologie: een zorg minder? Tilburg University. 2018. URL: https://www.tilburguniversity.edu/sites/default/files/download/Oratie_Wouters_TilburgU_5okt2018_1.pdf [accessed 2021-12-10]
30. Verkerk MJ, Holtkamp FC, Wouters EJ, van Hoof J. Professional practices and user practices: an explorative study in health care. Philos Reformata 2017 Dec;82(2):167-191. [doi: 10.1163/23528230-08202001]

31. Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, et al. Refining the theory of basic individual values. J Pers Soc Psychol 2012 Oct;103(4):663-688. [doi: 10.1037/a0029393] [Medline: 22823292]
32. Holtkamp FC, Wouters EJ, Verkerk MJ. Understanding user practices when drawing up requirements-the case of designing assistive devices for mobility. Int J Environ Res Public Health 2019 Jan 24;16(3):318 [FREE Full text] [doi: 10.3390/ijerph16030318] [Medline: 30678372]
33. Keller EJ, Crowley-Matoka M, Collins JD, Chrisman HB, Milad MP, Vogelzang RL. Fostering better policy adoption and inter-disciplinary communication in healthcare: a qualitative analysis of practicing physicians' common interests. PLoS One 2017 Feb 24;12(2):e0172865 [FREE Full text] [doi: 10.1371/journal.pone.0172865] [Medline: 28235088]
34. Kleyweg J. Moraliteitontwikkeling en de 10 waardentheorie van Schwartz: exploratief onderzoek naar de ontwikkeling van moreel oordelen in de maatschappij anno 2006 en de relatie tot de waardentheorie van Schwartz. Universiteit Twente. 2006 Aug 31. URL: https://essay.utwente.nl/58804/2/scriptie_J_Kleyweg.pdf [accessed 2023-08-12]
35. Fujihara R, Montezano L, Alfinito S. Are we ready for new technologies? The relation between human values and technology readiness applied to m-commerce in Brazil. Organ Soc 2022 Jul;29(102):537-568. [doi: 10.1590/1984-92302022v29n0024EN]
36. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med Ethics 2019 Dec 19;20(1):98 [FREE Full text] [doi: 10.1186/s12910-019-0437-z] [Medline: 31856798]
37. Dorsten AM, Sifford KS, Bharucha A, Mecca LP, Wactlar H. Ethical perspectives on emerging assistive technologies: insights from focus groups with stakeholders in long-term care facilities. J Empir Res Hum Res Ethics 2009 Mar;4(1):25-36. [doi: 10.1525/jer.2009.4.1.25] [Medline: 19374478]
38. Sallinen M, Hentonen O, Teeri S. Ethical dilemmas related to the use of safety technology in service house environments. Scand J Caring Sci 2020 Mar;34(1):199-205. [doi: 10.1111/scs.12721] [Medline: 31250937]
39. Rostad HM, Stokke R. Integrating welfare technology in long-term care services: nationwide cross-sectional survey study. J Med Internet Res 2021 Aug 16;23(8):e22316 [FREE Full text] [doi: 10.2196/22316] [Medline: 34398791]
40. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
41. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
42. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
43. Arets D, Ilievski D, Wernaart BF, Kamp JA, van Vliet M, Emini S. The right to Öffentlichkeit. In: Applied Human Rights. Wageningen, The Netherlands: Wageningen Academic; Jul 27, 2023.
44. Ponizovskiy V, Ardag M, Grigoryan L, Boyd R, Dobewall H, Holtz P. Development and validation of the personal values dictionary: a theory–driven tool for investigating references to basic human values in text. Eur J Pers 2020 Sep 01;34(5):885-902. [doi: 10.1002/per.2294]
45. Schwartz SH. An overview of the Schwartz theory of basic values. Online Read Psychol Cult 2012 Dec;2(1):1-20. [doi: 10.9707/2307-0919.1116]
46. Coahran M, Hillier LM, Van Bussel L, Black E, Churchyard R, Gutmanis I, et al. Automated fall detection technology in inpatient geriatric psychiatry: nurses' perceptions and lessons learned. Can J Aging 2018 Jul 03;37(3):245-260. [doi: 10.1017/s0714980818000181]
47. Nijhof N, van Gemert-Pijnen JE, de Jong GE, Ankoné JW, Seydel ER. How assistive technology can support dementia care: a study about the effects of the IST Vivago watch on patients' sleeping behavior and the care delivery process in a nursing home. Technol Disabil 2012 May 01;24(2):103-115. [doi: 10.3233/tad-2012-0339]
48. Niemeijer AR, Depla M, Frederiks B, Francke AL, Hertogh C. CE: original research: the use of surveillance technology in residential facilities for people with dementia or intellectual disabilities: a study among nurses and support staff. Am J Nurs 2014 Dec;114(12):28-37; quiz 38,48. [doi: 10.1097/01.NAJ.0000457408.38222.d0] [Medline: 25396755]
49. Holmes D, Teresi JA, Ramirez M, Ellis J, Eimicke J, Kong J, et al. An evaluation of a monitoring system intervention: falls, injuries, and affect in nursing homes. Clin Nurs Res 2007 Nov;16(4):317-335. [doi: 10.1177/1054773807307870] [Medline: 17991911]
50. Lauriks S, Meiland F, Osté JP, Hertogh C, Dröes RM. Effects of assistive home technology on quality of life and falls of people with dementia and job satisfaction of caregivers: results from a pilot randomized controlled trial. Assist Technol 2020 Sep 02;32(5):243-250 [FREE Full text] [doi: 10.1080/10400435.2018.1531952] [Medline: 30592439]
51. Dugstad J, Eide T, Nilsen ER, Eide H. Towards successful digital transformation through co-creation: a longitudinal study of a four-year implementation of digital monitoring technology in residential care for persons with dementia. BMC Health Serv Res 2019 Jun 10;19(1):366 [FREE Full text] [doi: 10.1186/s12913-019-4191-1] [Medline: 31182093]
52. Brodersen S, Lindegaard H. The smart floor: how a public-private partnership co-developed a heterogeneous healthcare technology system. Stud Health Technol Inform 2015;215:191-205. [Medline: 26249197]

53. Bourbonnais A, Rousseau J, Lalonde MH, Meunier J, Lapierre N, Gagnon MP. Conditions and ethical challenges that could influence the implementation of technologies in nursing homes: a qualitative study. Int J Older People Nurs 2019 Dec;14(4):e12266. [doi: 10.1111/opn.12266] [Medline: 31475466]
54. Berridge C, Halpern J, Levy K. Cameras on beds: the ethics of surveillance in nursing home rooms. AJOB Empir Bioeth 2019 Feb 22;10(1):55-62. [doi: 10.1080/23294515.2019.1568320] [Medline: 30794112]
55. Aloulou H, Mokhtari M, Tiberghien T, Biswas J, Phua C, Kenneth Lin JH, et al. Deployment of assistive living technology in a nursing home environment: methods and lessons learned. BMC Med Inform Decis Mak 2013 Apr 08;13(1):42 [FREE Full text] [doi: 10.1186/1472-6947-13-42] [Medline: 23565984]
56. Meng F, Song F, Guo M, Wang F, Feng X, Wang D, et al. Status and influential factors of intelligent healthcare in nursing homes in China. Comput Inform Nurs 2021 May 01;39(5):265-272 [FREE Full text] [doi: 10.1097/CIN.0000000000000685] [Medline: 33950900]
57. Sallinen M, Hentonen O, Kärki A. Technology and active agency of older adults living in service house environment. Disabil Rehabil Assist Technol 2015 Jan 17;10(1):27-31. [doi: 10.3109/17483107.2013.836685] [Medline: 24044649]
58. Fraile JA, Bajo J, Corchado JM, Abraham A. Applying wearable solutions in dependent environments. IEEE Trans Inform Technol Biomed 2010 Nov;14(6):1459-1467. [doi: 10.1109/titb.2010.2053849]
59. Bankole A, Anderson M, Knight A, Oh K, Smith-Jackson T, Hanson MA, et al. Continuous, non-invasive assessment of agitation in dementia using inertial body sensors. In: Proceedings of the 2nd Conference on Wireless Health. 2011 Presented at: WH '11; October 10-13, 2011; San Diego, CA. [doi: 10.1145/2077546.2077548]
60. Soto-Mendoza V, García-Macías JA, Chávez E, Martínez-García AI, Favela J, Serrano-Alvarado P, et al. Design of a predictive scheduling system to improve assisted living services for elders. ACM Trans Intell Syst Technol 2015 Jul 24;6(4):1-31 [FREE Full text] [doi: 10.1145/2736700]
61. Abbate S, Avvenuti M, Light J. Usability study of a wireless monitoring system among Alzheimer's disease elderly population. Int J Telemed Appl 2014;2014:617495-617498 [FREE Full text] [doi: 10.1155/2014/617495] [Medline: 24963289]
62. Anker-Hansen C, Grøndahl VA, Helgesen AK, Fagerli LB, Rummelhoff G, Bååth C, et al. Informal caregivers and assistive technology in Norwegian nursing homes. PLoS One 2022 Oct 4;17(10):e0275474 [FREE Full text] [doi: 10.1371/journal.pone.0275474] [Medline: 36194609]
63. Stark M, Tietz R, Gattinger H, Hantikainen V, Ott S. Effects of a mobility monitoring system on the cost of care in relation to reimbursement at Swiss nursing homes: learnings from a randomized controlled trial. Health Econ Rev 2017 Dec 01;7(1):43 [FREE Full text] [doi: 10.1186/s13561-017-0178-y] [Medline: 29192353]
64. Al‐Oraibi S, Fordham R, Lambert R. Impact and economic assessment of assistive technology in care homes in Norfolk, UK. J Assist Technol 2012;6(3):192-201. [doi: 10.1108/17549451211261317]
65. Engström M, Ljunggren B, Lindqvist R, Carlsson M. Staff perceptions of job satisfaction and life situation before and 6 and 12 months after increased information technology support in dementia care. J Telemed Telecare 2005;11(6):304-309. [doi: 10.1258/1357633054893292] [Medline: 16168167]
66. Offermann-van Heek J, Ziefle M. They don't care about us! Care personnel's perspectives on ambient assisted living technology usage: scenario-based survey study. JMIR Rehabil Assist Technol 2018 Sep 24;5(2):e10424 [FREE Full text] [doi: 10.2196/10424] [Medline: 30249592]
67. Te Boekhorst S, Depla MF, Francke AL, Twisk JW, Zwijsen SA, Hertogh CM. Quality of life of nursing-home residents with dementia subject to surveillance technology versus physical restraints: an explorative study. Int J Geriatr Psychiatry 2013 Apr;28(4):356-363. [doi: 10.1002/gps.3831] [Medline: 22644773]
68. Wigg JM. Liberating the wanderers: using technology to unlock doors for those living with dementia. Sociol Health Illn 2010 Feb 01;32(2):288-303 [FREE Full text] [doi: 10.1111/j.1467-9566.2009.01221.x] [Medline: 20422746]
69. Franco-Trigo L, Fernandez-Llimos F, Martínez-Martínez F, Benrimoj SI, Sabater-Hernández D. Stakeholder analysis in health innovation planning processes: a systematic scoping review. Health Policy 2020 Oct;124(10):1083-1099. [doi: 10.1016/j.healthpol.2020.06.012] [Medline: 32829927]
70. Riahi Y. Project stakeholders: analysis and management processes. Int J Econ Manag Stud 2017;4(3):39-45 [FREE Full text] [doi: 10.14445/23939125/IJEMS-V4I3P108]
71. Brugha R, Varvasovszky Z. Stakeholder analysis: a review. Health Policy Plan 2000 Sep;15(3):239-246. [doi: 10.1093/heapol/15.3.239] [Medline: 11012397]
72. Gregory R, Keeney RL. Creating policy alternatives using stakeholder values. Manag Sci 1994 Aug;40(8):1035-1048. [doi: 10.1287/mnsc.40.8.1035]
73. Braun D, Kramer J. Weerstand is bullshit. Boom Management. 2016 Oct 27. URL: https://boommanagement.nl/artikel/weerstand-is-bullshit/ [accessed 2024-01-14]
74. Lohss S. A new perspective on resistance to change. wibas. 2023 Feb 25. URL: https://www.wibas.com/blog/a-new-perspective-on-resistance-to-change/ [accessed 2024-01-14]
75. Willis R. What drives resistance to change? A leader's perspective. Assentire. 2012. URL: https://www.researchgate.net/publication/329514636_What_Drives_Resistance_to_Change_A_Leader's_Perspective [accessed 2024-01-12]

76. May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, et al. Understanding the implementation of complex interventions in health care: the normalization process model. BMC Health Serv Res 2007 Sep 19;7:148 [FREE Full text] [doi: 10.1186/1472-6963-7-148] [Medline: 17880693]
77. Braspenning AM, Cranen K, Snaphaan LJ, Wouters EJ. A multiple stakeholder perspective on the drivers and barriers for the implementation of lifestyle monitoring using infrared sensors to record movements for vulnerable older adults living alone at home: a qualitative study. Int J Environ Res Public Health 2022 Jan 05;19(1):570 [FREE Full text] [doi: 10.3390/ijerph19010570] [Medline: 35010829]
78. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
79. Verkerk M. A philosophy-based 'toolbox' for designing technology: the conceptual power of Dooyeweerdian philosophy. Koers- Bull Christ Scholarsh 2014 Aug 11;79(3). [doi: 10.4102/koers.v79i3.2164]
80. Dentoni D, Bitzer V, Schouten G. Harnessing wicked problems in multi-stakeholder partnerships. J Bus Ethics 2018 Apr 7;150:333-356. [doi: 10.1007/s10551-018-3858-6]
81. Oonk C, Gulikers J, den Brok P, Mulder M. Stimulating boundary crossing learning in a multi-stakeholder learning environment for sustainable development. Int J Sustain High Educ 2022 Jan 13;23(8):21-40. [doi: 10.1108/IJSHE-04-2021-0156]
82. Saidi T, Mork E, Aminoff S, Lobban F, Romm KL. Crossing boundaries in the delivery of healthcare - a qualitative study of an eHealth intervention in relation to boundary object theory. Digit Health 2023 Aug 30;9:20552076231196970 [FREE Full text] [doi: 10.1177/20552076231196970] [Medline: 37662674]
83. Santos Jhunior RO, Johnston NG, Boaventura JM, Barbero ER. Value co-creation within the stakeholder theory: taking stock and moving forward. Revista de Administração Unimep 2021 Jun;19(1) [FREE Full text]
84. Sorrentino M, Fiorilla C, Mercogliano M, Esposito F, Stilo I, Affinito G, et al. Technological interventions in European dementia care: a systematic review of acceptance and attitudes among people living with dementia, caregivers, and healthcare workers. Front Neurol 2024;15:1474336 [FREE Full text] [doi: 10.3389/fneur.2024.1474336] [Medline: 39416661]
85. Selderijk E. Aannames zijn killing voor je team. De Teamontwikkelaars. 2023 Mar 23. URL: https://deteamontwikkelaars.nl/aannames-zijn-killing-voor-je-team/ [accessed 2023-12-18]
86. Kloos N, Drossaert C, Bohlmeijer E, Westerhof GJ. How well do nursing staff assess the wellbeing of nursing home residents? An explorative study of using single-question scales. Ageing Soc 2020 Sep 22;42(4):812-828. [doi: 10.1017/S0144686X20001178]
87. Kunicki ZJ, Madrigal C, Quach LT, Riester MR, Jiang L, Duprey MS, et al. Comparing resident, proxy, and staff respondents for nursing home residents' preferences for everyday living. J Appl Gerontol 2023 Jan 26;42(1):28-36. [doi: 10.1177/07334648221123059] [Medline: 36029016]
88. Roelofs TS, Luijkx KG, Embregts PJ. A person-centered approach to study intimacy and sexuality in residential care facility (RCF) clients with dementia: methodological considerations and a guide to study design. NeuroRehabilitation 2017 Oct 14;41(2):359-373. [doi: 10.3233/NRE-172205] [Medline: 29060944]
89. van Boekel LC, Wouters EJ, Grimberg BM, van der Meer NJ, Luijkx KG. Perspectives of stakeholders on technology use in the care of community-living older adults with dementia: a systematic literature review. Healthcare (Basel) 2019 May 28;7(2):73 [FREE Full text] [doi: 10.3390/healthcare7020073] [Medline: 31141999]
90. Frank L, Jennings LA, Petersen RC, Majid T, Gilmore-Bykovskyi A, Schicker L, et al. Participation of persons with dementia and their caregivers in research. J Am Geriatr Soc 2021 Jul 10;69(7):1784-1792 [FREE Full text] [doi: 10.1111/jgs.17340] [Medline: 34245586]

Abbreviations

ICT: information and communication technology
LTC: long-term care
NPT: normalization process theory
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by R Bjarnadottir; submitted 11.07.24; peer-reviewed by J McMurray, M Nakanishi; comments to author 27.11.24; revised version received 09.01.25; accepted 30.01.25; published 20.03.25.

Please cite as:
van Gaans-Riteco D, Stoop A, Wouters E
Values of Stakeholders Involved in Applying Surveillance Technology for People With Dementia in Nursing Homes: Scoping Review
JMIR Aging 2025;8:e64074
URL: https://aging.jmir.org/2025/1/e64074
doi:10.2196/64074
PMID:39899267

Original Paper

Barriers to and Facilitators of Implementing Overnight Nursing Teleconsultation in Small, Rural Long-Term Care Facilities: Qualitative Interview Study

Veronique Nabelsi[1], BSc, MScA, PhD; Véronique Plouffe[2], BSc, MSc; Marie Chantal Leclerc[3], BSc, MSc, DBA

[1]Département des sciences administratives, Université du Québec, Québec, Canada

[2]Département des sciences comptables, Université du Québec en Outaouais, Gatineau, Canada

[3]Département des sciences infirmières, Université du Québec en Outaouais, Gatineau, Canada

Corresponding Author:
Veronique Nabelsi, BSc, MScA, PhD
Département des sciences administratives
Université du Québec
101, rue St-Jean-Bosco
Québec, J8X 3X7
Canada
Phone: 1 8195953900 ext 1915
Email: veronique.nabelsi@uqo.ca

Abstract

Background: Teleconsultation has expanded rapidly in recent years, especially during the COVID-19 pandemic, and has become standard practice among physicians. The benefits of teleconsultation, namely, improving access to care, ensuring continuity and quality of care, increasing patient satisfaction, and reducing costs and wait times, are well documented. However, its use in nursing practice, especially in long-term care settings, remains underresearched despite its significant transformative potential, particularly in resource-limited and rural settings, where it could address major challenges such as nursing shortages and access to care.

Objective: This study aimed to identify barriers to and facilitators of implementing overnight nursing teleconsultation in rural residential and long-term care centers in Quebec, Canada (*centres d'hébergement et de soins de longue durée* [CHSLDs]), with ≤50 beds.

Methods: A 6-month pilot project was rolled out sequentially in 3 rural CHSLDs in 2 administrative regions of Quebec between July 2022 and March 2023. A total of 38 semistructured interviews were conducted with managers (n=27, 71%), nursing staff members (n=9, 24%), and resident committee presidents (n=2, 5%) between February 2023 and July 2023.

Results: The study identified several barriers to the implementation of teleconsultation. The main barriers reported included union opposition (managers: 23/27, 85%), network instability (resident committee presidents: 2/2, 100%), limited technology skills (nursing staff members: 7/9, 78%), a perceived increase in workload (nursing staff members: 8/9, 89%; resident committee presidents: 2/2, 100%), and a low volume of teleconsultations (nursing staff members: 8/9, 89%). Despite the barriers, participants also identified key facilitators. These included the care setting (nursing staff members: 9/9, 100%; managers: 21/27, 78%), buy-in from senior management and managers (managers: 27/27, 100%; resident committee presidents: 2/2, 100%), collaboration between the departments (nursing staff members: 9/9, 100%), nursing staff motivation (nursing staff members: 9/9, 100%), and improvements in professional practices (nursing staff members: 8/9, 89%). Finally, the relative benefits of teleconsultation, such as enhanced mutual vision, faster assessment of clinical situations, improved resident care management quality, and greater flexibility and safety, were unanimously recognized (38/38, 100%) as contributing to its acceptability and potential for success.

Conclusions: This study provides an in-depth understanding of the barriers to and facilitators of implementing overnight nursing teleconsultation in small rural CHSLDs. This constitutes a sound basis for developing tailored strategies aimed at overcoming identified barriers and optimizing facilitators. The results also provide practical guidelines for decision makers, highlighting the need to adapt implementation approaches to the unique context of each facility. Furthermore, this study highlights the importance of further research to broaden our knowledge on the dissemination and scale-up of health care innovations. This includes the development of learning health systems capable of responding in an agile and effective way to the needs of rural and vulnerable populations both in Quebec and elsewhere.

(JMIR Aging 2025;8:e71950) doi:10.2196/71950

KEYWORDS

teleconsultation; long-term facilities; nursing; barriers and facilitators; rural; telehealth; qualitative; pilot study; Quebec

Introduction

Background

Teleconsultation has evolved considerably in recent years, especially during the COVID-19 pandemic, when it became standard practice for general practitioners and specialists alike [1-3]. Numerous studies have demonstrated its effectiveness in improving access to care [4-6], ensuring continuity [7,8] and quality of care [9,10], increasing patient satisfaction [11-15], and reducing costs [16-18] and wait times [2,19-22]. Moreover, internet-based clinical support initiatives between novice and expert professionals are being implemented in both urban [23-25] and rural [26,27] settings specifically to address challenges related to the shortage of qualified health care workers.

Despite the increasing adoption of teleconsultation, its use in nursing practice remains largely unexplored [28-31], even though its potential to transform care, including in long-term care centers, is well recognized [32,33]. This potential is even more significant in rural and resource-limited settings, where access to health care services remains a major challenge [34-37].

The lack of data on the use of teleconsultation in nursing practice is concerning as nurses play a key role in the continuum of care, especially in rural settings and long-term care facilities, where they are often patients' first point of contact [38-40]. Although recent studies have focused on the facilitators of and barriers to implementing geriatric teleconsultation in home nursing care [33,41], its impact on nurses' and nursing assistants' workflows [35], and the costs and cost savings associated with its use in residential and long-term care centers [34], these studies remain limited.

To date, studies have not yet provided a comprehensive picture of the effectiveness and impact of teleconsultation in nursing practice. Thus, in-depth exploration of this approach is essential to optimizing its benefits and supporting nursing professionals in the adoption and integration of this technological innovation.

Context

This research gap is of concern in residential and long-term care centers (*centres d'hébergement et de soins de longue durée* [CHSLDs]) in Quebec, Canada, where continuity and quality of care are essential to residents' well-being. Directives from the Ministry of Health and Social Services (Ministère de la Santé et des Services sociaux [MSSS]) and guidelines from the Order of Nurses of Quebec (Ordre des infirmières et infirmiers du Québec) stress the need to ensure the presence of nurses 24 hours a day in these facilities. However, smaller CHSLDs, namely, those with ≤50 beds, face major challenges in meeting this requirement due to nursing staff shortages, especially during night shifts. These frequent periods when no resources are available expose residents to increased risks to their safety and well-being [42,43]. This situation is even more critical in semiremote and remote areas, where difficulties associated with recruiting and retaining nursing staff exacerbate the challenges related to access and quality of care. A better understanding of the opportunities offered by teleconsultation could help alleviate these structural challenges and help build the CHSLDs' capacity in these settings.

As a response to these challenges, the National Directorate of Nursing Care and Services of the MSSS has initiated a pilot project to assess the impact of access to overnight nursing teleconsultation in rural CHSLDs with ≤50 beds. This initiative is based on the implementation of nursing teleconsultation, a promising solution to reinforce continuity of care and reduce regional disparities in access to health care services in Quebec.

Recent studies have shown that teleconsultation can mitigate limitations related to the lack of on-site nursing staff and offer real-time clinical support, thus reducing the risk of adverse events during periods of understaffing [35,44,45].

In line with this approach, this project explored innovative solutions to build resilience in long-term care systems and respond more effectively to the growing challenges associated with nursing shortages.

Objectives

The aim of this study was to identify barriers to and facilitators of implementing overnight nursing teleconsultation in rural Quebec CHSLDs with ≤50 beds. Specifically, this study aimed to gather the views of managers, nursing staff, and resident committee presidents. Exploring these viewpoints will fill a significant gap in the current literature and suggest critical avenues to support the successful integration of teleconsultation in long-term care settings.

Methods

Study Design and Setting

The 6-month pilot project was rolled out in 3 rural CHSLDs located in 2 administrative regions of Quebec. The regions were selected by the MSSS for their alignment with the project's outlined criteria, which include facilities located in semiremote and remote areas, those already experiencing nursing shortages during the night shift, and those reporting issues and risks related to these shortages. In addition, at least 30% of all CHSLD facilities in the territory have a capacity of ≤50 beds. The rollout was conducted sequentially from July 2022 to March 2023 at different sites. Given the innovative nature of the pilot project, an exploratory qualitative study was conducted to identify the barriers to and facilitators of implementing teleconsultation in overnight nursing care.

Data Collection

An interview guide was designed, tested, and validated by the research team. The guide, comprising 12 open-ended questions, aimed to identify the barriers to and facilitators of implementing

teleconsultation. This guide provided a better understanding of the context and experiences surrounding the pilot project's deployment. This study was guided by key factors influencing the implementation of health innovations as outlined in the framework proposed by Chaudoir et al [46]. The interview guide is included in Multimedia Appendix 1.

The framework developed by Chaudoir et al [46] was selected for several key reasons that ensure that it is an appropriate framework to assess the implementation of health innovations. In fact, this model is based on a systematic review of the literature. It also captures the complexity of implementing health innovations by considering various levels of influence, such as organizational structures, health care providers, patients, and the specific characteristics of the innovation.

This holistic approach enables a more comprehensive and nuanced assessment of the factors that can affect an innovation's success. Furthermore, by incorporating levels of analysis that reflect practical realities in the field—such as the nursing staff and residents—the model provides health care professionals with an applicable and relevant framework. It helps identify the barriers and facilitators that are specific to each setting, thus facilitating the design of customized implementation strategies. Finally, the model is flexible and can be tailored to different health innovations and settings. Whether we are examining the implementation of a new technology or a care protocol, the model proposes analysis categories that can be adjusted according to the features of the innovation and environment.

The model's components can be described on five levels: (1) the structural level, which includes factors related to the broader context in which the innovation is implemented, such as health care policies, funding structures, or available resources. These variables directly or indirectly influence an organization's ability to adopt and integrate new practices. (2) the organizational level, which focuses on the specific characteristics of the organizations themselves, such as organizational culture, leadership, support systems, and communication dynamics. It examines how these internal elements facilitate or hinder the implementation of innovations. (3) the health care provider level; at this level, the focus is on the individuals responsible for implementation, such as physicians, nurses, or other health care professionals. The model assesses health care providers' attitudes, knowledge, skills, and beliefs, all of which can influence how an innovation is adopted and applied. (4) the patient level, which analyzes patients' perceptions, attitudes, and behaviors, as well as their level of innovation engagement and buy-in. Patients' psychosocial factors, such as their understanding, beliefs, or preferences, are crucial to the successful implementation of health innovations. (5) the innovation level, which examines the specific characteristics of the innovation itself, such as complexity, compatibility with existing practices, cost, and flexibility. An innovation perceived as easy to use, relevant, and beneficial is more likely to be adopted.

This framework provides a solid foundation to analyze the perspectives of managers, nursing staff, and resident committee presidents, highlighting the key factors influencing the adoption and effectiveness of nursing teleconsultation in small CHSLDs.

Participants

Participant recruitment was conducted using nonprobability sampling [47], through which participants were identified by pilot project managers in each region. This selection method is designed to maximize participants' intrinsic motivation by giving them the opportunity to become involved on their own terms. By promoting this freedom of choice, this study aimed to attract participants who were especially motivated and engaged, thereby improving the quality of the collected data as well as enhancing the relevance and validity of the findings.

Participants were initially contacted through an email that included the interview guide and a detailed consent form. The consent form outlined the context of the study, the project's objectives, the procedures and expected duration of participation, the anticipated benefits, and assurances regarding anonymity and confidentiality. A total of 38 semistructured individual interviews were conducted with managers (n=27, 71%), nursing staff (n=9, 24%), and resident committee presidents (n=2, 5%) between February 2023 and July 2023.

These semistructured interviews were carried out in French and were conducted via videoconference (Zoom; Zoom Video Communications, Inc). The principal investigator (VN) conducted the interviews. Participants were given the option to review the transcripts of each interview, but none of the participants chose to receive the transcripts. No additional recruitment process was necessary as information redundancy indicated data saturation [48].

Data Analysis

On the basis of the framework by Chaudoir et al [46], we conducted a pattern analysis of interview transcripts to identify factors describing facilitators of and barriers to implementing overnight nursing teleconsultation in small CHSLDs. This approach emphasized hierarchical coding, enabling rigorous structuring of the textual data analysis while offering the flexibility required to meet the specific needs of the study.

The analysis prioritized participants' responses, highlighting their descriptions of barriers and facilitators. To do this, we immersed ourselves in the data by reading and rereading transcripts while taking handwritten notes on emerging factors and codes. This iterative process carried out by a single coder (the principal investigator, VN) is a valid method in qualitative analysis of thematic data, enabling researchers to understand how participants gave meaning to their experiences. A handwritten thematic map was created to group data extracts into broad categories of barriers and facilitators, which fostered a thorough review process and helped generate initial codes. These codes were then applied during a second analysis phase using the NVivo software (version 14; QSR International). The initial codes were used to identify overlaps and search for emerging themes.

To ensure a comprehensive and structured analysis, the coder (VN) applied the 5 levels of the framework by Chaudoir et al [46]—structural, organizational, health care provider, patient, and innovation—as primary coding categories. Each identified barrier or facilitator was initially coded under one of these levels. The coder developed subthemes within each level, enabling a

more nuanced examination of the data. This structured approach reinforced the consistency of our analysis, ensuring that all aspects of the implementation process were thoroughly explored while maintaining alignment with established theoretical constructs.

Finally, potential factors were examined and verified across the dataset by rereading the transcripts and checking themes against identified codes. This approach ensured the robustness and consistency of the findings.

Ethical Considerations

Ethics approval was obtained from the Research Ethics Committee of the Outaouais Integrated Health and Social Services Centre before the beginning of the study (reference 2022-353_195) in Quebec, Canada. All participants gave their consent electronically before beginning the interviews. Participation was anonymous and voluntary. Study participants did not receive monetary compensation. All interviews were audio recorded with the participants' permission and then transcribed in compliance with ethical and confidentiality standards. The deidentified recordings were transcribed verbatim by a third-party transcription service bound by a confidentiality agreement. The study's findings will be disseminated through presentations at conferences and publications in peer-reviewed journals using anonymized data. The findings will also be shared through presentations to various MSSS stakeholders and the nursing community.

Results

Overview of Barriers

This study aimed to identify the barriers to and facilitators of implementing overnight nursing teleconsultation in 3 of Quebec's rural CHSLDs with ≤50 beds. The results are presented in accordance with the 5 levels of the framework by Chaudoir et al [46]: structural, organizational, health care provider, patient, and innovation.

Table 1 presents a comprehensive overview of the framework detailing the barriers identified at each level and for each participant group.

Table 1. Level, factors, and number of participants who mentioned each factor.

Level and factor	Managers[a] (n=27), n (%)	Nursing staff members[b] (n=9), n (%)	Resident committee presidents[c] (n=2), n (%)
Structural			
Union opposition	23 (85)	—[d]	—
Network instability	11 (41)	4 (44)	2 (100)
Overburdened managers	16 (59)	—	—
Lack of support from project leaders	—	6 (67)	—
Organizational			
Lack of leadership from the site manager	—	4 (44)	—
Health care provider			
Resistance to change	12 (44)	6 (67)	—
Limited technology skills	12 (44)	7 (78)	—
Insecurity about using technology	12 (44)	4 (44)	—
Increased workload associated with the technology	10 (37)	8 (89)	2 (100)
Patient			
Concerns about quality of care	—	—	1 (50)
Innovation			
Low volume of teleconsultations	14 (52)	8 (89)	—
Complexity of the process compared to a phone call with a remote on-call nurse	16 (59)	—	—
Increased time to initiate care management	—	5 (56)	—
Insecurity about the quality of nursing assessments	—	5 (56)	—
Difficulty using the tablet	—	4 (44)	—

[a]Individuals in positions of authority. They oversee operations, manage resources, and supervise personnel in health care settings. Their role is to ensure the smooth functioning of the facility.

[b]This refers to all individuals involved in providing direct care to patients or residents, including both nurses and nursing assistants.

[c]These individuals lead the residents' committee. The committee represents residents in health care facilities such as long-term care homes. Its mandate is to protect residents' rights. The committee ensures that residents are treated with dignity and that their rights and freedoms are respected. It also serves as a key spokesperson for residents. It brings residents' concerns and needs to the attention of the institution's governing bodies.

[d]Not applicable.

Barriers: Structural-Level Factors

Union Opposition

Union opposition was mentioned by most managers (23/27, 85%) as a major factor before the launch of the pilot project. According to respondents' testimonials, for several months, the union waged a disinformation campaign conveying alarming messages to health care professionals and the community.

The union disseminated messages stating that the "government was planning to replace nurses with tablets," which it claimed would "diminish the quality of care and endanger the safety of residents" (M10). Another manager (M9) specified that "They ran a lot of ads saying nurses were being replaced by tablets." Some managers reported that this campaign had a disruptive effect, creating "a shock wave and a wave of fear in the community" (M22), and the pilot project was perceived "as devaluing this client group, as if they were receiving second-class care" (M25).

The union's intervention was not confined to the public sphere; it also manifested itself directly in the workplace. Managers reported that union representatives visited CHSLDs attempting to dissuade nurses from participating in the pilot project. One manager (M8) explained the following:

> *The union would come directly into the workplace, to frighten employees.*

The union told nursing assistants that their participation was "super dangerous, because they would be going beyond their scope of practice and their professional order would turn against them," added one manager (M9). Despite this initial pressure, none of the interviewed nursing staff members (9/9, 100%) reported any problems with their union after the beginning of the pilot project. Unlike the managers, nursing staff members did not perceive union opposition as a barrier.

Network Instability

Network instability was identified as a barrier by some managers (11/27, 41%) and nursing staff members (4/9, 44%), as well as

by the resident committee presidents (2/2, 100%), especially in rural and remote areas.

Interviewees' testimonials revealed that connectivity was not uniform across the facilities. One manager (M1) explained the following:

> *It didn't necessarily work everywhere in the CHSLD.*

In total, 44% (4/9) of nursing staff members added that there were sometimes 15- to 20-minute delays in logging in. This wasted time, although occasional, can have significant repercussions on the quality of care, as mentioned by one resident committee president (RC36):

> *From time to time, it won't work there...you have to take that into account, because it would be a huge waste of time to start an intervention, then you lose the network, you have to restart another way, by telephone, etc.*

Overburdened Managers

The extra workload associated with the pilot project was a barrier for 59% (16/27) of the managers, especially for project leaders, project coleaders, and site managers. They had to reconcile their usual tasks while ensuring rapid deployment of the project within the context of a nursing shortage. One manager (M4) explained the following:

> *We weren't optimal in our monitoring, which created an obstacle, because, basically, we weren't as present.*

Moreover, daily monitoring of nursing staff practices increased the burden on site managers, who feared that implementation of the project would not be feasible without additional resources. One manager (M9) raised the following question:

> *Is the workload going to be realistic for managers who are already highly solicited?*

Lack of Support From Project Leaders

Lack of support from project leaders was identified as a barrier by 67% (6/9) of nursing staff members. This limited support manifested as reduced availability as project leaders were often overwhelmed by their many responsibilities. One nursing staff member (NS34) indicated the following:

> *We'd have liked to have a little more time, but they're kind of busy with everything.*

This created a feeling of abandonment among nursing staff, with some expressing a lack of guidance. One staff member (NS29) reported that "The project leaders didn't always have the answers to our questions."

Barriers: Organizational-Level Factors (Lack of Leadership From the Site Manager)

In one administrative region, the site manager's lack of leadership was perceived as a barrier by some nursing staff members (4/9, 44%). This lack of leadership manifested itself as a lack of proximity between the site manager and the nursing staff, namely, infrequent travel to meet the team and gaps in communication. This distance hindered the exchange of information and the understanding of the pilot project's issues, leading the site manager to become disinterested, which generated frustration and reduced nursing staff buy-in and motivation at the outset of the pilot project. However, as the project progressed, the situation improved, and they were able to overcome these challenges.

Barriers: Health Care Provider–Level Factors

Resistance to Change

Resistance to change was identified as a barrier by 44% (12/27) of managers and 67% (6/9) of nursing staff members. This resistance took the form of a reluctance to participate in training and simulations, as well as a marked preference for using the telephone, which was perceived as quicker and more effective. For example, one nurse refused to use teleconsultation to assess a resident after a fall, preferring to travel to the CHSLD herself. One manager (M3) indicated the following:

> *The nursing staff found it cumbersome.*

Indeed, nursing staff members reported that their colleagues preferred traditional methods such as on-call nursing or traveling to visit the resident in person. Even when teleconsultation would have been more appropriate, using the telephone remained the preferred method.

Limited Technology Skills

Limited technology skills represented a barrier, as identified by 67% (18/27) of the managers. One manager (M20) pointed out the following:

> *One of the barriers we encountered very, very quickly was that people were not familiar with the technology and then were not able to use it.*

According to one manager (M23), this shortcoming can be explained by "a lack of simulation and comfort as well as by constraints such as nursing shortages, heavy workloads, and COVID-19 outbreaks." One manager (M8) said the following:

> *In addition to learning new tools, our nurses had to continue providing care. At one point, they were saying, "It's just not working".... It was a real challenge to implement this on a daily basis.*

According to 78% (7/9) of nursing staff members, this barrier is especially significant among older nurses, who are often less comfortable with new technology. One nursing staff member (NS33) described this generational challenge:

> *We weren't all born with a keyboard or tablet in our hands...we have a few who are in their fifties. Not all of them were comfortable with it either.*

One nursing staff member (NS29) added the following:

> *...although the younger generation showed an initial interest, this desire was curbed by older nurses' reluctance to embrace the technology, creating a barrier to the successful integration of digital tools into professional practices.*

Furthermore, the low volume of remote activities and insufficient monitoring limited the practice of teleconsultation, resulting in the loss of acquired skills. One nursing staff member (NS33) described it as follows:

> *We had practices, and I fell on the practice that was postponed.... It was never rescheduled. So, the first time, I'd never even practiced.*

Insecurity About Using Technology

Insecurity about using technology was a concern due to the novelty of teleconsultation and was mentioned by 44% (12/27) of managers. In addition, 44% (4/9) of care staff members reported that nursing assistants in particular felt vulnerable when they were alone on-site, fearing that they would not know how to use teleconsultation properly or solve technical problems. As one manager (M4) explained, "It's my nurse who takes charge" when a resident is not doing well. Another manager (M9) confirmed this feeling of insecurity:

> *This practice, in terms of the technology, well, it caused a little insecurity at first, because we didn't know that much about it.*

Similarly, one nursing staff member (NS32) noted that "a fear of computer technology" discouraged some of the more experienced nurses from becoming involved in the pilot project.

Increased Workload Associated With the Technology

The implementation of teleconsultation led to an increased workload for remote nurses, a challenge that was highlighted by 37% (10/27) of managers, 89% (8/9) of nursing staff members, and 100% (2/2) of resident committee presidents. One nursing staff member (NS31) reported the following:

> *We were the ones who had to adapt the most.*

Before every night shift, nurses had to make sure that they had the necessary tools for teleconsultation, which often meant traveling to the CHSLD even on their days off. In addition to their usual tasks, they had to manage interdepartmental reports, fill in specific follow-up forms, and immediately document each teleconsultation. One resident committee president (RC36) explained the following:

> *We also had to foresee working time to connect, use the equipment, get everything working and then provide electronic notes afterward, transfer them, etc.*

One manager (M23) pointed out the following:

> *...the people who are working remotely, the ones who are doing the teleconsultation, are the same people who are there in the evening, during the night...it creates a great deal of anxiety.*

Barriers: Patient-Level Factors (Concerns About Quality of Care)

In the context of teleconsultation, concerns about the quality of care were primarily raised by 50% (1/2) of the resident committee presidents. He feared that teleconsultation would undermine the personal nature of care, concerned that technology would compromise the human contact that is essential to in-person interactions. Initially opposed to the project, fearing the impersonality and disempowerment of nurses, he eventually recognized the benefits of teleconsultation as the pilot progressed. One of the presidents (RC36) expressed the following:

> *My initial opposition to the project was based on what I didn't want: That it would become impersonal, that it would prevent human contact...that it would change the on-call nurse's responsibility, relying on a screen, which is not the same as what you might experience during an in-person visit.*

Barriers: Innovation-Level Factors

Low Volume of Teleconsultations

According to 52% (14/27) of managers, the low volume of teleconsultations was a barrier and could be attributed to the small size of CHSLDs, but the low volume of activity also raised important questions, as one manager pointed out (M25):

> *Was the volume low because there wasn't a need? Was the volume low because practices were already good on both sides, and [the person carrying out an intervention] acted preventively?*

Another manager (M4) added the following:

> *Nursing assistants who said, "Oh no, look, it's 4 o'clock. The nurse is coming in two hours, we'll wait two hours." Did this harm the resident? Well, indirectly, for someone who is in pain, yes it did. But there was no report of an incident or accident that could have or did cause harm to the resident's health, safety, and well-being.*

According to 89% (8/9) of nursing staff members, the low volume of teleconsultations hindered their ability to maintain their skills. One of them (NS26) said the following:

> *It had been a month since we'd had one...I forgot to fill in the smartsheet.*

Moreover, a manager (M10) observed the following:

> *...when comparing data from the previous year to [the data related to] the implementation of teleconsultation, the number of telephone calls received is equal to the number of teleconsultations over the same period.*

Complexity of the Teleconsultation Process Compared to On-Call Nursing

The complexity of the teleconsultation process compared to on-call nursing was perceived as a barrier by 59% (16/27) of managers. Unlike on-call nursing, when the nursing assistant can contact the on-call nurse directly, teleconsultation involves a series of more complex actions, such as waking up the on-call nurse to initiate the consultation and, sometimes, the need to call back the CHSLD. This complexity prolongs the time it takes to obtain a nursing assessment, as explained by one manager (M18):

> *It's too slow...the time to turn on the laptop, to connect safely.*

Another manager (M10) added that the speed of the on-call process influenced perceptions of teleconsultation:

> *It influenced the teleconsultation project.*

In addition, the necessity of having the teleconsultation on-call travel case added another layer of difficulty, especially when staff forgot the equipment, as described by one manager (M12):

> *Ah OK, but now I don't have the equipment, I have to go to the CHSLD to get the equipment.*

Increased Time to Initiate Care Management

According to 56% (5/9) of nursing staff members, the use of teleconsultation led to an increase in the time taken to initiate the residents' care management. Unlike previous practices, the nurse had to assess the resident over a digital platform before intervening, adding an extra step that delayed the response to immediate needs. This delay was exacerbated in remote areas, where unstable internet connections complicated access to teleconsultation, potentially leading to a deterioration of the resident's condition. One nursing staff member (NS26) illustrated this problem by describing a situation in which the requirement to use a tablet for teleconsultation interfered with the prompt management of a resident's pain:

> *A lady was experiencing pain.... I found the computer-based support detrimental to immediate care.... Meanwhile, the lady was in pain. You know, we are managing pain at the same time as we manage the tablet.*

Insecurity About the Quality of Nursing Assessments

There were concerns about the quality of nursing assessments carried out via teleconsultation, including the fear that visual assessment cannot adequately replace a physical examination. In total, 56% (5/9) of nursing staff members shared these concerns. One of them (NS30) explained the following:

> *I had concerns about the physical assessment in the sense that shifting to a visual assessment instead of doing it in real life...that my assessment would not be complete.*

The absence of physical contact with the resident was perceived as a limitation as nonverbal communication plays an important part in a comprehensive assessment. Another nursing staff member (NS33) illustrated this difficulty:

> *Non-verbal and verbal [messages] contradict each other in residents...with the tablet, it wasn't easy because I couldn't look at my resident's face and leg movement at the same time.*

The use of technology such as the digital stethoscope also prompted reservations. One nurse expressed unease:

> *Listening over the phone, it's not like performing the auscultation myself.* [NS31]

Difficulty Using the Tablet

When using a tablet, nursing staff members experienced physical constraints in terms of mobility and effectiveness. These constraints made it difficult to carry out teleconsultations, as one nurse described:

> *The hindrance was caused by the darn arm they set up to hold that tablet.... It would swing around. You know, to be honest, it wasn't the best.* [NS33]

A total of 44% (4/9) of the nurses emphasized the need for a stand to hold the tablet, freeing up the nursing assistant's hands.

Overview of Facilitators

Table 2 presents a comprehensive overview of the framework detailing the facilitators identified at each level and for each participant group.

Table 2. Level, dimension, and number and percentage of participants who mentioned each factor.

Level and dimension	Managers[a] (n=27), n (%)	Nursing staff members[b] (n=9), n (%)	Resident committee presidents[c] (n=2), n (%)
Structural			
Care setting	21 (78)	9 (100)	—[d]
Legitimization of the practice of overnight on-call nursing	18 (67)	—	—
Culture of on-call nursing	—	—	2 (100)
Implementation monitoring	21 (78)	—	—
Organizational			
Buy-in from senior management and managers	27 (100)	—	2 (100)
Support from project leaders	19 (70)	—	—
Support from site managers	19 (70)	7 (78)	—
Team involvement, motivation, and stability	15 (56)	—	—
Collaboration between the nursing department and the Support Program for the Autonomy of Seniors	18 (67)	9 (100)	—
Transfer of knowledge and experience	21 (78)	—	—
Health care provider			
Nursing staff buy-in	13 (48)	—	—
Nursing staff motivation	—	9 (100)	—
Development of the nursing staff's skills	13 (48)	—	—
Ability to adapt and use technology	—	7 (78)	—
Openness to change	—	—	2 (100)
Patient			
Buy-in from residents, families, and resident committees	12 (44)	—	—
Communication	—	6 (67)	—
Innovation			
Relative benefits	27 (100)	9 (100)	2 (100)
Development of nursing staff's roles	19 (70)	—	—
Improved professional practices	—	8 (89)	—

[a]Individuals in positions of authority. They oversee operations, manage resources, and supervise personnel in health care settings. Their role is to ensure the smooth functioning of the facility.

[b]This refers to all individuals involved in providing direct care to patients or residents, including both nurses and nursing assistants.

[c]These individuals lead the residents' committee. The committee represents residents in health care facilities such as long-term care homes. Its mandate is to protect residents' rights. The committee ensures that residents are treated with dignity and that their rights and freedoms are respected. It also serves as a key spokesperson for residents. It brings residents' concerns and needs to the attention of the institution's governing bodies.

[d]Not applicable.

Facilitators: Structural-Level Factors

Care Setting

Testimonials from 78% (21/27) of the managers and all nursing staff members (9/9, 100%) highlighted the care setting's decisive role in the success of the pilot project. Faced with a nursing staff shortage, this initiative was viewed as a promising solution to optimize practices and the management of available resources while maintaining the quality and safety of resident care. One manager (M1) explained the following:

> *Our objective is to ensure that every nursing staff member is in the right place, playing their role to the full and that we are using our resources wisely.*

Legitimization of the Practice of Overnight On-Call Nursing

Legitimization of the practice of overnight on-call nursing factored positively in the project's success. According to 67% (18/27) of managers, the fact that this practice was framed within a specific, temporary context reassured stakeholders such as the Order of Nursing Assistants of Quebec (Ordre des infirmières

et infirmiers auxiliaires du Québec), the Quebec ombudsman, and user and resident committees. One manager (M1) explained the following:

> *This legitimization enabled nursing staff to feel they had greater authorization to use teleconsultation, mitigating fears related to professional compliance.*

Culture of On-Call Nursing

The existing culture of on-call nursing was a significant facilitator. All resident committee presidents (2/2, 100%) stated that this culture, which was already well established and accepted in CHSLDs, facilitated the implementation of teleconsultation. Considered "an innovative and adaptive solution, on-call nursing was seen as essential to maintaining optimal quality of care," as indicated by one of the presidents (RC36). He added the following:

> *...the pre-existing culture facilitated the transition to teleconsultation by normalizing the idea of a remote nurse and positioning it as a safe and effective approach.*

Implementation Monitoring

Close monitoring by the nursing department (ND) and the Support Program for the Autonomy of Seniors (SAPA) within the health care system in Quebec was a key facilitator according to 78% (21/27) of managers. Monitoring took place at three levels: (1) strategic level—committees and regular meetings with the MSSS promoted fluid communication on project advancement; (2) operational level—project leaders organized regular meetings with site managers, enabling actions to be adjusted quickly and providing immediate feedback (as one manager [M16] reported, "monitoring by the Ministry...was highly beneficial"); and (3) day-to-day level—ongoing monitoring of teleconsultation practice was implemented, including a review of ministry forms, hospital transfers, and incident reports, as well as audits in CHSLDs to ensure that nursing staff had the support they needed.

Facilitators: Organizational-Level Factors

Buy-In From Senior Management and Managers

Senior management buy-in was viewed as a facilitator by all managers (27/27, 100%), namely owing to the support of the chief executive officer of the Integrated Health and Social Services Centre of Abitibi-Témiscamingue and the Integrated University Health and Social Services Centre of Mauricie and Centre-du-Québec, as reported by one manager (M8):

> *He took it on, then he defended it.*

When senior management prioritizes a project, it motivates other managers to engage, facilitating rollout and the resolution of challenges such as acquiring equipment—"We received our equipment very quickly because it was a priority," according to one manager (M1). The resident committee presidents (2/2, 100%) also confirmed that this support was important to the project's success. One of the resident committee presidents (RC35) explained the following:

> *When senior management is supportive of the project, it sends a strong signal to the members of the organization about the strategic importance of teleconsultation. This approval from leadership can positively influence the levels of acceptance and engagement within the team.*

In addition, the managers' buy-in was unanimously recognized as a facilitator by participating managers (27/27, 100%). Their engagement made it possible to effectively navigate MSSS requirements and ensure the buy-in of the nursing staff members who were consulted, underlining the importance of creating an environment that is conducive to the implementation of teleconsultation.

Support From Project Leaders

Support from project leaders was a determining factor according to 70% (19/27) of managers. Project leaders played a key role in motivating nursing staff by clarifying objectives, allaying concerns, and fostering champions within the teams, creating a conducive environment for the adoption of teleconsultation. As one manager (M5) explained, "Having a dedicated person to answer questions and solve problems" was essential.

Support From Site Managers

According to 70% (19/27) of managers, site managers also fostered the implementation of teleconsultation. One manager (M12) reported the following:

> *Their knowledge of the environment and their existing bonds of trust played a key role in the human management of change and in nursing staff mobilization.*

The nursing staff (7/9, 78%) also appreciated this support, underlining the site managers' guidance and availability. One nursing staff member (NS32) explained the following:

> *I felt supported throughout the project. If I had any questions, I knew where to turn. I had a lot of support from my manager.*

Involvement, Motivation, and Stability of Nursing Staff

According to 56% (15/27) of managers, the involvement, motivation, and stability of nursing staff were key facilitators. Team cohesion facilitated flexibility and mutual support, as stated by one manager (M16):

> *Nursing staff demonstrated solidarity by swapping shifts during snowstorms to ensure staff availability.*

This solidarity enabled staff to respond effectively to residents' needs and maintain reasonable response times.

Collaboration Between the ND and SAPA

According to 67% (18/27) of managers, collaboration between the ND and SAPA was a key facilitator. A clear division of roles enabled the ND to manage external relationships with the MSSS and other agencies, whereas the SAPA dealt directly with the implementation of teleconsultation, ensuring effective coordination. One manager (M16) reported the following:

> *Having a single point of entry was helpful.*

This synergy promoted the cocreation of solutions to the project's challenges, namely in terms of training and monitoring, reinforcing the effectiveness and success of the initiative.

The training plan, including coaching, simulations, and tool adaptations, was tailored to meet regional needs and reinforce the safety of nursing practices. One manager (M10) underlined the following:

> *Training was customized...to reinforce safety.*

Practical guides and equipment such as headsets supported the practice of teleconsultation, and simulations boosted the nursing staff's confidence. Nursing staff members (9/9, 100%) unanimously appreciated the training, deeming it essential to the adoption of the technology and success of the pilot project.

Transfer of Knowledge and Experience

The transfer of knowledge and experience was a significant factor for most managers (21/27, 78%), facilitating collaboration and adaptation between regions and within participating CHSLDs.

Between regions, managers worked closely, sharing their experiences and adjusting approaches according to the specific needs of each region. Although support tools were not cocreated systematically, these exchanges enabled participants to adjust based on local context. Regarding collaboration and adaptation within CHSLDs, in one region, the level of sharing between 2 CHSLDs was particularly striking. Nursing staff from the first site shared their experiences with that of the second, fostering buy-in to the pilot project. For example, a manager's guide created from lessons learned was passed on to the other site, facilitating the implementation of teleconsultation in similar settings. One manager (M16) described it as follows:

> *The manager took notes on everything she had implemented...and brought it back to the ND....So, it ranged from the criteria we had to meet, to making sure we met them, to the tasks we had to carry out...all in one guide.*

Facilitators: Health Care Provider–Level Factors

Nursing Staff Buy-In

For 48% (13/27) of managers, nursing staff buy-in was a facilitator. Nursing staff members were not only favorable to the idea, they were also motivated to actively participate in the pilot project, demonstrating a willingness to move forward with teleconsultation. One manager (M5) noted the following:

> *The nursing staff were very open and aligned with the project. They wanted to move forward with the change.*

Despite initial stress, the nursing staff adapted quickly. As one manager (M24) pointed out, "Once the adaptation period was over...there was no more stress. It went well," underlining their ability to overcome resistance and make a successful transition to teleconsultation.

Nursing Staff Motivation

All nursing staff members (9/9, 100%) considered that the motivation of health care staff was a facilitator. The main sources of motivation included commitment to the team, interest in technological tools, and the desire to help resolve the nursing shortage. Nursing staff members appreciated the creation of overnight nursing assistant positions with remote support, helping address the shortage and improve the quality of care. As one staff member put it, "We're not as effective after 1 p.m" (NS28), highlighting the challenges of working long hours. Being motivated to use technological tools such as teleconsultation reflects a desire to explore innovative solutions to improve working conditions and better meet residents' needs.

Development of the Nursing Staff's Skills

Development of the nursing staff's skills was a key factor, as highlighted by 48% (13/27) of managers. Training tailored to the staff's needs fostered their preparedness and engagement, making them champions of the pilot project. Younger members showed greater mastery of computer and technological skills. In addition, the training improved nursing assistants' level of autonomy and clinical judgment, contributing to the development of clinical leadership and team management skills. One manager (M7) described it as follows:

> *When she's on her own, she also takes on a bit of a coordination role...something she didn't usually do because it was part of the nurse's role.*

Ability to Adapt and Use Technology

For 78% (7/9) of nursing staff members, the ability to use technology was directly related to computer literacy and adaptability. Despite a low level of initial fluency, some managed to overcome their difficulties. One staff member (NS33) put it as follows:

> *I'm not very tech-savvy. I managed. It went well.*

Thus, adaptability was a key factor, with nursing staff members demonstrating an ability to adjust to teleconsultation, including those who were not as comfortable with computers.

Openness to Change

Nursing staff's openness to change was essential according to both resident committee presidents (2/2, 100%). This positive attitude was fostered by the fact that the project was recognized as an opportunity for exploration. One of the resident committee presidents (RC36) said the following:

> *Given that it was a pilot project, everyone agreed to give it a try.*

RC35 added the following:

> *...the transfer of experience from nurses performing on-call nursing to teleconsultation is also a concrete example of this openness to change. The fact that the staff had prior experience in similar practices made it easier to adapt to new methods.*

Facilitators: Resident-Level Factors

Buy-In From Residents, Families, and Resident Committees

Buy-in from residents, families, and resident committees greatly facilitated the implementation of teleconsultation according to 44% (12/27) of managers. This support was reinforced by the creation of a relationship of trust through transparent

communication and regular meetings with the site manager. One manager (M17) pointed out the following:

> *The site manager kept us informed on a regular basis, establishing a climate of trust.*

Communication

Communication was a facilitator, especially in interactions with residents and their families. In total, 67% (6/9) of nursing staff members noted that proactive communication with residents and their families, namely, explaining the project, answering questions, and obtaining informed consent, facilitated resident buy-in. One nursing staff member (NS30) said the following:

> *Transparency, especially about concerns such as data leaks, fostered a positive reception to teleconsultation.*

Families generally welcomed the technology, recognizing the additional benefit to care delivery. Nursing staff members observed that there was no negative impact on residents, which could be attributed to the effective communication that reassured families about data confidentiality and security.

Facilitators: Innovation-Level Factors

Relative Benefits

Analysis of the relative benefits of teleconsultation showed that this modality represented a significant added value for the dyad composed of the remote nurse and the CHSLD nursing assistant according to all managers (27/27, 100%), nursing staff members (9/9, 100%), and resident committee presidents (2/2, 100%). Benefits included, first, improved mutual vision. All nursing staff members (9/9, 100%) agreed that teleconsultation enabled direct observation of the resident and of the nursing assistant's nonverbal cues, fostering a greater understanding of the situation than was possible through on-call nursing. One staff member (NS31) pointed out the following:

> *The way [the nursing assistants] report it to us over the phone and the way we see it through our own assessment, are two different things. Sometimes it's minimized, and sometimes it's exaggerated.*

This visual component enhanced the remote nurse's ability to carry out a more accurate and thorough assessment. One of the resident committee presidents (RC35) stated the following:

> *...the use of video in teleconsultation provides a clear advantage in terms of assessment quality compared to voice-only interactions. This ability to visualize the patient can lead to more informed decisions regarding necessary interventions.*

The second benefit was faster assessment. Teleconsultation reduced the wait time for nursing assessments, enabling more effective interventions according to all nursing staff members (9/9, 100%). One staff member (NS32) noted the following:

> *...an intervention that might have taken an hour could take only 15 minutes with teleconsultation.*

Moreover, the direct electronic transmission of nursing notes to the CHSLD promoted continuity of care. One of the resident committee presidents (RC36) noted the following:

> *...that despite the implementation of teleconsultation, it is not intended to completely replace in-person visits, but rather to complement care, thus offering flexibility and adaptability in the delivery of healthcare services.*

The third benefit was improved quality of care management. The visual component of teleconsultation helped identify signs or symptoms that were not described verbally, leading to more informed decisions according to all nursing staff members (9/9, 100%). In addition, one manager (M25) reported that teleconsultation "provides visual support that the telephone does not, enabling [the nurse] to validate the on-site nursing assistant's hypothesis or to support her in her contribution to the assessment."

The fourth benefit was flexibility and safety. According to the resident committee presidents (2/2, 100%), teleconsultation did not replace in-person visits but was complementary to them, providing flexibility in the delivery of health care services. One of the presidents (RC36) mentioned the following:

> *Nurses are not prevented from physically traveling if necessary.*

This innovation also contributed to resident safety according to the other president (RC35), noting that "residents were safe with this technological innovation, reinforcing the idea that teleconsultation does not entail any compromise in terms of patient safety."

All managers (27/27, 100%) reported that the introduction of teleconsultation was perceived as a major step forward, enhancing the quality and safety of care delivery. According to them, teleconsultation provided greater safety by enabling nurses to exercise their clinical judgment under improved conditions. One manager (M4) stated the following:

> *Teleconsultation gives the manager a sense of security because they know that the healthcare professional is in a better position to exercise their clinical judgment.*

Development of Nursing Staff's Roles

The pilot project contributed significantly to the development of nursing staff's roles according to 70% (19/27) of managers. Although overnight on-call nursing was already integrated into the organizational culture, the pilot project improved the organization and quality of care by strengthening nursing staff's roles.

Key improvements included enhancement of the remote nurse and nursing assistant dyad as the nursing assistant's practice was strengthened, including more effective nursing assessments and better care planning during the night shift, and expanded scope of practice for nursing assistants as changes to nursing care procedures enabled nursing assistants to fully exercise their skills. One manager (M1) pointed out the following:

> *Nursing assistants have activities...that we've allowed them to carry out.*

Another manager (M16) added that the pilot project enabled them to "apply their entire scope of practice, enhancing the value and recognition of their role."

Improved Professional Practices

The implementation of teleconsultation significantly improved professional nursing practices according to most nursing staff members (8/9, 89%). The main observed benefits were, first, the development of nursing assistants' skills. Teleconsultation enabled nursing assistants to develop their scope of practice, acquiring new skills and playing a more active role in assessing and delivering care. Structuring communication when transmitting information also helped strengthen their communication skills. The second main observed benefit was proactive information updates. Teleconsultation facilitated the updating of therapeutic nursing plans and medication administration records, enhancing nursing assistants' autonomy. One nursing staff member explained that teleconsultation eliminated the need to constantly contact the nurse for simple decisions such as administering medication, ensuring greater autonomy for nursing assistants. This staff member (NS29) underlined that "this project has really helped to make us more autonomous...we can manage almost everything on our own."

Discussion

Principal Findings

Overview

The aim of this study was to gain a better understanding of the experiences of managers, nursing staff, and resident committee presidents involved in the pilot project to identify the factors that may promote the implementation of teleconsultation. Our multilevel analysis revealed the presence of facilitators and barriers.

Structural-Level Factors

Among the identified structural barriers, union opposition, which was reported by most managers (23/27, 85%), initially represented a significant obstacle, but it subsided once the project was in place. The union recognized that teleconsultation complied with professional standards while at the same time complementing the work done by nurses in CHSLDs without seeking to replace it. This helped dispel any initial fears. In addition, managers' leadership played a decisive role by supporting the nursing staff, responding to their concerns, and facilitating a smooth and harmonious transition.

However, network instability in rural areas was a barrier for some managers (11/27, 41%) and nursing staff members (4/9, 44%). Connectivity issues coupled with the lack of overnight technical support led to delays and interruptions, compromising the effectiveness and reliability of teleconsultation. These difficulties highlighted the limitations of technological infrastructures in these environments and created situations in which interventions were delayed, compromising continuity of care. This problem is corroborated by studies that underline the importance of stable network connectivity to ensure the smooth operation of teleconsultation services. Indeed, network quality is a determining factor in ensuring the fluidity of exchanges and the effectiveness of remote consultations, as highlighted by several research studies [28,49-52]. In these settings, ensuring reliable network coverage and efficient technical support is imperative to avoid interruptions that could adversely affect the quality of care.

Overburdened managers, who had to juggle their usual responsibilities with the demands of the pilot project, led to suboptimal management according to 59% (16/27) of managers. This double workload, compounded by a nursing shortage, made it difficult to adequately monitor the project and led to management being less present and reactive. This finding is in line with observations found in the literature, which stress that work overload is a major challenge in project management, especially in settings with limited resources [53]. Thus, managers' inability to respond optimally to project requirements due to the pressure of their day-to-day responsibilities contributed to the suboptimal implementation of the initiative, underlining the need for greater support for managers during the implementation of complex projects.

Lack of support from project leaders was a barrier for 67% (6/9) of nursing staff members. Insufficient training and monitoring limited the nursing staff's ability to use teleconsultation effectively. This situation reflects the importance of organizational support, which studies have shown to be a key factor in the success of training and implementation programs [54]. Indeed, constant support and adequate coaching help build the staff's skills and ensure the successful adoption of new technology, such as teleconsultation, in care settings.

However, several facilitators contributed to the project's success. The context of the health care system, marked by a nursing shortage that was exacerbated by the COVID-19 pandemic, acted as a catalyst for the implementation of teleconsultation according to several managers (21/27, 78%) and all nursing staff members (9/9, 100%). Perceived as an effective solution to the shortage, teleconsultation enabled better management of human resources, which reduced excessively long shifts and improved the nursing staff's working conditions. This approach is supported by studies demonstrating that technology can optimize the use of human resources during a shortage [55,56].

Some managers (18/27, 67%) highlighted that legitimization of the practice of overnight on-call nursing played a key role in nursing staff's acceptance and adoption of teleconsultation. According to the literature, framing new clinical practices within a defined, transparent, and temporary framework is essential to reassuring stakeholders such as professional orders, user committees, and regulatory bodies [35,57,58]. Such a framework helps build an environment of trust, reducing reluctance to adopt innovative practices and ensuring compliance with professional standards.

The preexisting culture of on-call nursing facilitated the acceptance of the pilot project by making the transition to teleconsultation smoother according to the resident committee presidents (2/2, 100%). Familiarity with remote practices fostered acceptance of the new technology. This observation is supported by the literature, which indicates that preexisting practices and organizational cultures play a decisive role in the acceptance of health care innovations [35,59].

Finally, close monitoring at several levels was a key factor in the successful implementation of teleconsultation for a large proportion of managers (21/27, 78%). By integrating strategic, operational, and daily monitoring, the project benefited from rapid adjustments, ensuring compliance with project objectives and requirements. This systematic monitoring not only facilitated proactive management of challenges but also enhanced responsiveness to emerging issues, helping maintain coherent and fluid processes. According to the literature, structured and sustained monitoring is essential to optimizing the implementation of new practices, especially in a context of technological transformation, helping overcome barriers and ensure project sustainability [60,61].

Organizational-Level Factors

According to the nursing staff in 1 of the 2 administrative regions (4/9, 44%), one of the barriers at the organizational level was the site managers' lack of leadership. This lack of leadership, characterized by minimal presence and limited communication with nursing staff, was viewed as a barrier to the implementation of teleconsultation. The perceived distance between the site managers and care teams created a climate of frustration and disengagement. This situation is corroborated by the literature, which underlines that weak leadership can hinder the implementation of change initiatives by generating feelings of frustration and disengagement among health care teams [62-64].

As part of the implementation of teleconsultation, several organizational factors facilitated its rollout, highlighting the importance of a coordinated approach and sustained engagement at all levels of the organization. Buy-in and active support from senior management and managers were crucial to the project's success according to all managers (27/27, 100%) and resident committee presidents (2/2, 100%). The literature supports this observation, stating that the engagement of organizational leaders is a key factor in the success of change initiatives in health care facilities [65,66]. This engagement helped mobilize the necessary resources and promote a culture conducive to innovation.

The role of project leaders and site managers was considered essential by 70% (19/27) of managers. Their availability and expertise not only enabled the effective resolution of operational issues but also played a key role in maintaining the nursing staff's level of motivation.

Specifically, nursing staff members (7/9, 78%) stated that regular meetings and personalized support helped clarify objectives, allayed concerns, and ensured constant monitoring, which helped overcome challenges and optimize implementation processes. These results are in line with the literature, which emphasizes the importance of management practices and managers' commitment to the success of digital transformation initiatives [67,68].

According to some managers (15/27, 56%), the involvement, motivation, and stability of nursing staff members were also facilitators. The staff's ability to adapt and maintain a high level of service despite challenges was facilitated by increased motivation and cohesiveness, which is supported by the work by Nabelsi et al [35]. Their research shows that staff motivation and stability are essential elements to ensuring efficient processes in health care. The solidarity and cooperation observed within the teams helped maintain high levels of service even under difficult conditions.

Another facilitator was the collaboration between the ND and the SAPA according to all nursing staff members (9/9, 100%) and several managers (18/27, 67%). This collaboration clarified the division of roles, avoided duplication, and ensured smooth project management. A clear division of responsibilities and coordination between departments are essential [35]. This model of cross-directorate collaboration helped maximize efficiency and avoid overlaps that might slow down the implementation process.

Finally, most managers (21/27, 78%) said that the transfer of knowledge and experience between regions and within CHSLDs played a significant role in localizing the project. This approach, which centered on collaboration and the sharing of best practices, enhanced the effectiveness of the pilot project [35]. The study by Nabelsi et al [35] shows that knowledge sharing between teams and sites not only fosters the adoption of technology, it also enables better management of the challenges encountered in the field by adapting to local needs and constraints.

Health Care Provider–Level Factors

Resistance to change was a barrier to the implementation of teleconsultation, especially for certain nursing staff members (6/9, 67%). Their marked preference for the telephone, perceived as faster and more effective, highlighted their resistance to the adoption of teleconsultation. The literature on organizational change in health care indicates that this resistance may be fueled by the perception of being overloaded and by deeply ingrained habits, making it difficult to accept new technology [28,69,70].

According to nursing staff members (7/9, 78%), a lack of skill in using new technology was also a limiting factor, especially among older nursing staff members, who were not as comfortable with technological tools. This shortcoming is a recognized factor in the failure to implement technology in health care [44,71,72]. Work overload, exacerbated by crises such as the COVID-19 pandemic, added a further dimension to this challenge, creating a cognitive overload that complicated the effective integration of new technology.

Insecurity about using technology was another barrier for a small proportion of managers (12/27, 44%) and nursing staff members (4/9, 44%). Nursing assistants in particular felt vulnerable due to their lack of familiarity with teleconsultation and fear of dealing with technical problems. The literature on technology acceptance underlines that insecurity and fear of the unknown can reduce user motivation and performance [73,74].

The increased workload associated with technology was also a barrier. The implementation of teleconsultation led to increased management of reports, specific forms, and detailed documentation, an issue recognized by most nursing staff members (8/9, 89%) as well as all resident committee presidents (2/2, 100%). Studies show that increased workloads can cause

stress and reduce job satisfaction, negatively influencing the implementation of new practices [75,76].

Among the facilitators, nursing staff buy-in was important, although it was mentioned by a lower proportion of managers (13/27, 48%). The nursing staff's willingness to participate in the pilot project and their positive attitude toward teleconsultation were key to its success. The literature on organizational change in health care suggests that stakeholder buy-in is important to the success of change initiatives [77,78].

Nursing staff members' motivation also played a significant role in the pilot project's success, a factor unanimously recognized by all staff members (9/9, 100%). Commitment to the team, interest in technological tools, and the desire to help solve the nursing shortage were all motivating factors. Research shows that motivation is a facilitator of the acceptance and successful use of health technology [79,80].

Development of the nursing staff's skills was a facilitator, although it was mentioned by a lower proportion of managers (13/27, 48%). Adequate training strengthened nurses' readiness and commitment, enabling them to become project champions. The literature indicates that skill development is key to the successful adoption of new technology [66,81,82]. Customized training and support tools such as practical guides and simulations helped build the nursing staff's skills and confidence [35].

The ability to use the technology was also crucial to the success of the pilot project, as underlined by nursing staff members (7/9, 78%). Rapid adaptation to new technology is indeed a facilitator, as confirmed in research on technology acceptance [74,80].

Finally, openness to change facilitated the implementation of the project, a factor that was unanimously recognized by resident committee presidents (2/2, 100%). This openness to change led to smoother adoption of teleconsultation, which is supported by studies demonstrating that acceptance of change is essential to the success of transformation initiatives [83].

Resident-Level Factors

The only barrier was fear concerning the quality of care provided via teleconsultation, which was mentioned by 50% (1/2) of the resident committee presidents. One of the presidents expressed concerns about the reduction in human contact, which could lead to perceived depersonalization of care. However, these concerns were dispelled as the pilot progressed, and he eventually recognized the benefits of teleconsultation. This fear that technology would disempower nurses and create a sense of impersonality represented a barrier to the acceptance of teleconsultation. Studies indicate that concerns about quality of care and depersonalization can negatively influence acceptance of telehealth technology by residents and their families [50,72,84,85].

According to some managers (12/27, 44%), buy-in from residents, families, and resident committees was a facilitator of the implementation of teleconsultation. The managers observed that transparent communication and regular meetings with these stakeholders helped establish a climate of trust. This approach is in line with the literature, which stresses the importance of trust and effective communication when fostering acceptance of health technology by patients and their families [86-88].

For most nursing staff members (6/9, 67%), effective communication itself was a facilitator. Research shows that managing expectations and clarifying the benefits of new technology are important to their acceptance by residents and their families. The ability to clearly explain the benefits of teleconsultation and address residents' concerns contributed greatly to their buy-in [35].

Innovation-Level Factors

The low volume of teleconsultations was recognized as a barrier by more than half (14/27, 52%) of the managers. Several possible explanations were put forward. Some managers suggested that this low volume may reflect a lack of real need, the preexisting effectiveness of preventive practices, or some nursing staff members' reluctance to use teleconsultation. However, when comparing data from the previous year with the data related to implementation of teleconsultation, findings reveal that the number of telephone calls received was equal to the number of teleconsultations over the same period. In addition, the low volume of consultations impacted the nursing staff's ability to maintain their skills. Most nursing staff members (8/9, 89%) reported a decrease in their level of comfort with the technological tools due to sporadic use.

The complexity of the teleconsultation process was also a barrier for more than half (16/27, 59%) of the managers. Compared to traditional on-call nursing, teleconsultation involves more complex technological processes. Connolly et al [89] underline that this complexity can reduce the effectiveness of interventions and increase staff frustration, hindering the adoption and effectiveness of technology [51,90].

Another barrier was the increased time to initiate care management according to over half (5/9, 56%) of nursing staff members. Using digital platforms for assessments can increase response times, a problem that is exacerbated by technological limitations and connectivity issues. The research by Pilosof et al [91] shows that these delays can adversely impact the quality of care by affecting the responsiveness of interventions.

Concerns about the quality of assessments conducted via teleconsultation were noteworthy. Just over half (5/9, 56%) of nursing staff members expressed concern about the ability of visual assessment to effectively replace a physical assessment, highlighting the potential risk of compromising quality of care. This fear is corroborated by studies revealing that telemedicine can sometimes alter the quality of clinical assessments if not properly integrated into care practices [85-87,92-95].

Finally, the difficulty of using a tablet for teleconsultations represented a barrier for a small proportion of nursing staff members (4/9, 44%). Srinivasan et al [85] highlight that ergonomic issues and difficulties in handling technological equipment can reduce the effectiveness of interventions and user satisfaction, complicating the integration of teleconsultation [86,87,89].

Despite these challenges, several innovation-level factors facilitated the implementation of teleconsultation. The relative

benefits of this technology were viewed positively by all managers (27/27, 100%), nursing staff members (9/9, 100%), and resident committee presidents (2/2, 100%). They appreciated the improved mutual vision and faster assessment. Teleconsultation solutions offer significant benefits in terms of visual communication and speed of intervention [35]. This perception of the benefits fostered acceptability and support for the project.

The development of nursing staff's roles was another facilitator according to several managers (19/27, 70%). The project strengthened professional practices and broadened the nursing assistants' skills. The work by Nabelsi et al [35] indicates that teleconsultation technology can support the expansion of professional roles and improve quality of care.

Moreover, the implementation of teleconsultation led to improved professional practices for most nursing staff members (8/9, 89%). Research has shown that technology integration improves professional skills and care management, highlighting the potential of innovations to positively transform practices in care settings [35,91,96].

Limitations and Future Research

This study has a number of limitations that must be taken into account when interpreting the results. First, this study was conducted in only 2 Quebec regions, including a total of 3 CHSLDs. This limited scope may restrict the generalizability of the results to other geographical settings or types of facilities. In addition, the small number of sites included in the study may not enable researchers to capture the diversity of practices and challenges encountered in other regions or in facilities of different sizes. Second, this study focused exclusively on smaller facilities with ≤50 beds. While this is in line with the study's objective of targeting small CHSLDs, the results may not be directly applicable to larger facilities, which may have different organizational structures and needs. Finally, this study's evaluation period was short, making it impossible to observe the long-term impact of nursing teleconsultation, especially in terms of the continuous improvement of nursing practices, the sustainability of interventions, and the changes in stakeholder perceptions.

Furthermore, this study used a nonprobabilistic sampling method, which may have resulted in the inclusion of participants who were more inclined to view teleconsultation favorably or who had a particular interest in the topic. To mitigate this potential selection bias, the team researcher actively sought a diversity of perspectives during data collection, encouraging participants to share both supportive and critical viewpoints regarding the implementation of teleconsultation. Moreover, a rigorous qualitative analysis was conducted, with particular attention given to dissenting opinions and negative experiences, ensuring a comprehensive representation of the facilitators and barriers encountered. Despite these efforts, the inherent limitations of qualitative research, including the subjectivity of self-reported experiences, necessitate further investigation through complementary methodologies. A quantitative study conducted through a survey would strengthen the robustness of our findings and allow for a more generalizable assessment of the impact of teleconsultation

To broaden our understanding of the implementation of nursing teleconsultation in long-term care, it would be relevant to conduct studies in a larger number of CHSLDs, including facilities of various sizes located in different regions, to assess the transferability of this study's findings and their effectiveness on a larger scale. Longitudinal research would also be needed to assess the long-term effects of nursing teleconsultation on quality of care, resident satisfaction, and human resource management in CHSLDs, as well as to identify any adjustments needed to ensure the sustainability of these practices.

It would also be useful to study the impact of teleconsultation on nurses' well-being and workload, namely by examining how this technology can be optimized to effectively support their role without increasing their stress level or mental load. Finally, economic studies could help quantify the costs associated with implementing nursing teleconsultation and compare them with potential savings in terms of decreased hospitalizations, adverse events, and improved quality of care.

Conclusions

This study provides the first in-depth analysis of barriers and facilitators related to the implementation of overnight nursing teleconsultation in small long-term care facilities. The findings provide a better understanding of these barriers, which can be used to develop strategies to overcome them during implementation. These findings are also particularly relevant to decision makers who are responsible for designing health initiatives as their choices influence the implementation and scaling-up process.

Broadly, the results provide a comprehensive overview of the factors influencing the successful implementation of teleconsultation in long-term care. This can help identify key factors to consider when scaling up teleconsultation in CHSLDs. The framework developed by Chaudoir et al [46] highlights the concept of adaptability, emphasizing the importance of adjusting the deployment of an innovation to suit the specific context. When scaling up teleconsultation, it is important to consider the specific characteristics of each CHSLD and region and tailor the implementation of teleconsultation accordingly.

While resistance to change is often considered a major barrier to implementing new health care technologies, our findings challenge this assumption. In the rural CHSLDs studied, the preexisting on-call nursing culture not only facilitated the adoption of teleconsultation but also eased its integration into clinical practice. This contrasts with previous research suggesting that health care professionals may resist new technology due to concerns about workflow disruptions or unfamiliarity with remote care models. In this context, previous experience with remote support likely mitigated these challenges, highlighting the importance of accounting for contextual factors when implementing teleconsultation. Furthermore, the identification of a low volume of teleconsultations as a barrier contradicts the common assumption that a phased rollout is always beneficial. Instead, our results suggest that achieving a critical mass of teleconsultations is essential to maintaining staff engagement and competencies. These findings provide new insights into teleconsultation implementation by demonstrating how preexisting practices

and use patterns can significantly influence the adoption and sustainability of technological innovations in long-term care settings.

Efforts to implement overnight nursing teleconsultation in long-term care are more likely to succeed if they are based on an understanding of the forces driving the dissemination and scale-up of teleconsultation. Therefore, further research is needed to develop and strengthen the conceptual and applied foundations of the dissemination and scale-up of health care innovations, especially in the context of Quebec's emerging learning health care systems.

Acknowledgments

This study was funded and supported by the Ministry of Health and Social Services (Ministère de la Santé et des Services sociaux) in Quebec, Canada. The authors are most grateful to the health professionals and managers of the Integrated Health and Social Services Centre of Abitibi-Témiscamingue and the Integrated University Health and Social Services Centre of Mauricie and Centre-du-Québec. The authors wish to thank everyone who participated in interviews and discussions.

Authors' Contributions

VN conceptualized and designed the study, collected the data, conducted the interviews, conducted the analysis, and wrote the first draft of the manuscript. VN and VP have read and revised the manuscript. All authors (VN, VP, and MCL) approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 16 KB - aging_v8i1e71950_app1.docx]

References

1. Fernández Coves A, Yeung KH, van der Putten IM, Nelson EA. Teleconsultation adoption since COVID-19: comparison of barriers and facilitators in primary care settings in Hong Kong and the Netherlands. Health Policy 2022 Oct;126(10):933-944 [FREE Full text] [doi: 10.1016/j.healthpol.2022.07.012] [Medline: 36050194]
2. Smith AC, Thomas E, Snoswell CL, Haydon H, Mehrotra A, Clemensen J, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Jun;26(5):309-313 [FREE Full text] [doi: 10.1177/1357633X20916567] [Medline: 32196391]
3. Omboni S, Padwal RS, Alessa T, Benczúr B, Green BB, Hubbard I, et al. The worldwide impact of telemedicine during COVID-19: current evidence and recommendations for the future. Connect Health 2022 Jan 04;1:7-35 [FREE Full text] [doi: 10.20517/ch.2021.03] [Medline: 35233563]
4. Williams S, Barnard A, Collis P, Correia de Sousa J, Ghimire S, Habib M, et al. Remote consultations in primary care across low-, middle- and high-income countries: implications for policy and care delivery. J Health Serv Res Policy 2023 Jul;28(3):181-189 [FREE Full text] [doi: 10.1177/13558196221140318] [Medline: 36484225]
5. Wherton J, Greenhalgh T, Hughes G, Shaw SE. The role of information infrastructures in scaling up video consultations during COVID-19: mixed methods case study into opportunity, disruption, and exposure. J Med Internet Res 2022 Nov 10;24(11):e42431 [FREE Full text] [doi: 10.2196/42431] [Medline: 36282978]
6. Salisbury C, Murphy M, Duncan P. The impact of digital-first consultations on workload in general practice: modeling study. J Med Internet Res 2020 Jun 16;22(6):e18203 [FREE Full text] [doi: 10.2196/18203] [Medline: 32543441]
7. Norberg BL, Austad B, Kristiansen E, Zanaboni P, Getz LO. The impact and wider implications of remote consultations for general practice in Norway: qualitative study among Norwegian contract general practitioners. JMIR Form Res 2024 Dec 17;8:e63068 [FREE Full text] [doi: 10.2196/63068] [Medline: 39688890]
8. Xuan Y, Guo C, Lu W. The effects of information continuity and interpersonal continuity on physician services online: cross-sectional study. JMIR Med Inform 2022 Jul 21;10(7):e35830 [FREE Full text] [doi: 10.2196/35830] [Medline: 35862183]
9. Hawrysz L, Kludacz-Alessandri M, Walczak R. Predictive factors of physicians' satisfaction and quality of work under teleconsultation conditions: structural equation analysis. JMIR Hum Factors 2024 Jun 10;11:e47810 [FREE Full text] [doi: 10.2196/47810] [Medline: 38857081]
10. Campbell K, Greenfield G, Li E, O'Brien N, Hayhoe B, Beaney T, et al. The impact of virtual consultations on the quality of primary care: systematic review. J Med Internet Res 2023 Aug 30;25:e48920 [FREE Full text] [doi: 10.2196/48920] [Medline: 37647117]

11. Meng G, McAiney C, McKillop I, Perlman CM, Tsao S, Chen H. Factors that influence patient satisfaction with the service quality of home-based teleconsultation during the COVID-19 pandemic: cross-sectional survey study. JMIR Cardio 2024 Feb 16;8:e51439 [FREE Full text] [doi: 10.2196/51439] [Medline: 38363590]
12. Andrews E, Berghofer K, Long J, Prescott A, Caboral-Stevens M. Corrigendum to "satisfaction with use of telehealth during COVID-19: an integrative review" [international journal of nursing studies advanced 2 (2020). Article number 100008]. Int J Nurs Stud Adv 2022 Dec;4:100092 [FREE Full text] [doi: 10.1016/j.ijnsa.2022.100092] [Medline: 35989885]
13. Hong YR, Lawrence J, Williams Jr D, Mainous III A. Population-level interest and telehealth capacity of us hospitals in response to COVID-19: cross-sectional analysis of google search and national hospital survey data. JMIR Public Health Surveill 2020 Apr 07;6(2):e18961 [FREE Full text] [doi: 10.2196/18961] [Medline: 32250963]
14. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
15. Tenforde AS, Hefner JE, Kodish-Wachs JE, Iaccarino MA, Paganoni S. Telehealth in physical medicine and rehabilitation: a narrative review. PM R 2017 May;9(5S):S51-S58. [doi: 10.1016/j.pmrj.2017.02.013] [Medline: 28527504]
16. Lin Y, Xu X, Liu Y, Alias H, Hu Z, Wong LP. Perception and acceptance of telemedicine use in health care among the general public in China: web-based cross-sectional survey. J Med Internet Res 2024 Jul 16;26:e53497 [FREE Full text] [doi: 10.2196/53497] [Medline: 39012687]
17. Huguet M, Sarazin M, Perrier L, Augusto V. How we can reap the full benefit of teleconsultations: economic evaluation combined with a performance evaluation through a discrete-event simulation. J Med Internet Res 2022 May 20;24(5):e32002 [FREE Full text] [doi: 10.2196/32002] [Medline: 35594065]
18. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298 [FREE Full text] [doi: 10.2196/17298] [Medline: 33074157]
19. Nabelsi V, Lévesque-Chouinard A. Successful electronic consultation service initiative in Quebec, Canada with primary care physicians' and specialists' experiences on acceptance and use of technological innovation: cross-sectional exploratory study. JMIR Form Res 2024 May 30;8:e52921 [FREE Full text] [doi: 10.2196/52921] [Medline: 38814689]
20. Almathami HK, Win KT, Vlahu-Gjorgievska E. Empirical evidence of internal and external factors influencing users' motivation toward teleconsultation use. Telemed J E Health 2024 Jan;30(1):141-156. [doi: 10.1089/tmj.2022.0527] [Medline: 37343179]
21. Kichloo A, Albosta M, Dettloff K, Wani F, El-Amir Z, Singh J, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530 [FREE Full text] [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
22. Nabelsi V, Lévesque-Chouinard A, Liddy C, Dumas Pilon M. Improving the referral process, timeliness, effectiveness, and equity of access to specialist medical services through electronic consultation: pilot study. JMIR Med Inform 2019 Jul 10;7(3):e13354 [FREE Full text] [doi: 10.2196/13354] [Medline: 31293239]
23. Roberson AE, Carlson M, Kohler CM, Harris PA, Volkmann CL. Initiating virtual nursing in general inpatient care. Am J Nurs 2023 Jun 01;123(6):48-54. [doi: 10.1097/01.NAJ.0000938736.42266.5e] [Medline: 37233140]
24. Savitz ST, Frederick RK, Sangaralingham LR, Lampman MA, Anderson SS, Habermann EB, et al. Evaluation of safety and care outcomes after the introduction of a virtual registered nurse model. Health Serv Res 2023 Oct;58(5):999-1013. [doi: 10.1111/1475-6773.14208] [Medline: 37525521]
25. Steaban RL, Morrison J, Simmons S, Ivory C, France D, Leming-Lee S, et al. Strategies to evaluate new models of nursing care to meet hospital staffing and patient care needs. Nurse Lead 2024 Dec;22(6):718-724. [doi: 10.1016/j.mnl.2024.05.007]
26. Nott S, Wingfield G, Haigh A, Luscombe GM, Thompson AE, Saurman E, et al. The Virtual Rural Generalist Service: a hybrid virtual model of care designed to improve health access and outcomes in rural and remote communities. Med J Aust 2024 Dec 09;221 Suppl 11:S3-S7. [doi: 10.5694/mja2.52529] [Medline: 39647931]
27. Thompson AE, Saurman E, Nott S, Wilson A, Shaw T. Clinician experiences of a hybrid virtual medical service supporting rural and remote hospitals: a qualitative study. Med J Aust 2024 Dec 09;221 Suppl 11:S16-S21. [doi: 10.5694/mja2.52525] [Medline: 39647921]
28. Regragui S, Abou Malham S, Gaboury I, Bois C, Deville-Stoetzel N, Maillet L, et al. Nursing practice and teleconsultations in a pandemic context: a mixed-methods study. J Clin Nurs 2023 Sep;32(17-18):6339-6353. [doi: 10.1111/jocn.16756] [Medline: 37202866]
29. Vaughan C, Lukewich J, Mathews M, Hedden L, Poitras ME, Asghari S, et al. Nursing contributions to virtual models of care in primary care: a scoping review protocol. BMJ Open 2022 Sep 20;12(9):e065779 [FREE Full text] [doi: 10.1136/bmjopen-2022-065779] [Medline: 36127080]
30. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021 Jun 14;373:n1190. [doi: 10.1136/bmj.n1190]
31. Bashshur R, Doarn CR, Frenk JM, Kvedar JC, Woolliscroft JO. Telemedicine and the COVID-19 pandemic, lessons for the future. Telemed J E Health 2020 May;26(5):571-573. [doi: 10.1089/tmj.2020.29040.rb] [Medline: 32275485]

32. Dai Z. Telehealth in long-term care facilities during the COVID-19 pandemic - lessons learned from patients, physicians, nurses and healthcare workers. BMC Digit Health 2023;1(1):2 [FREE Full text] [doi: 10.1186/s44247-022-00003-y] [Medline: 38014371]
33. Tan AJ, Rusli KD, McKenna L, Tan LL, Liaw SY. Telemedicine experiences and perspectives of healthcare providers in long-term care: a scoping review. J Telemed Telecare 2024 Feb;30(2):230-249. [doi: 10.1177/1357633X211049206] [Medline: 34666535]
34. Nabelsi V, Plouffe V. Assessing cost and cost savings of teleconsultation in long-term care facilities: a time-driven activity-based costing analysis within a value-based healthcare framework. BMC Health Serv Res 2024 Sep 13;24(1):1064 [FREE Full text] [doi: 10.1186/s12913-024-11578-9] [Medline: 39272121]
35. Nabelsi V, Leclerc MC, Plouffe V. Nurses' and nursing assistants' experiences with teleconsultation in small rural long-term care facilities: semistructured interview pilot study. JMIR Aging 2024 Nov 27;7:e65111 [FREE Full text] [doi: 10.2196/65111] [Medline: 39602791]
36. Özkaytan Y, Schulz-Nieswandt F, Zank S. Acute health care provision in rural long-term care facilities: a scoping review of integrated care models. J Am Med Dir Assoc 2023 Oct;24(10):1447-57.e1. [doi: 10.1016/j.jamda.2023.06.013] [Medline: 37488029]
37. Qiao Y, Ran L, Li J. Optimization of teleconsultation using discrete-event simulation from a data-driven perspective. Telemed J E Health 2020 Jan 01;26(1):114-125. [doi: 10.1089/tmj.2018.0229] [Medline: 30994407]
38. Abdolkhani R, Petersen S, Walter R, Zhao L, Butler-Henderson K, Livesay K. The impact of digital health transformation driven by COVID-19 on nursing practice: systematic literature review. JMIR Nurs 2022 Aug 30;5(1):e40348 [FREE Full text] [doi: 10.2196/40348] [Medline: 35867838]
39. Santos MT, Halberstadt BM, Trindade CR, Lima MA, Aued GK. Continuity and coordination of care: conceptual interface and nurses' contributions. Rev Esc Enferm USP 2022;56:e20220100 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2022-0100en] [Medline: 35951450]
40. Cacchione PZ. Innovative care models across settings: providing nursing care to older adults. Geriatr Nurs 2020;41(1):16-20. [doi: 10.1016/j.gerinurse.2020.01.011] [Medline: 32033809]
41. Qi Tan AJ, Chua WL, McKenna L, Chin Tan LL, Lim YJ, Liaw SY. Enablers and barriers to nurse-facilitated geriatric teleconsultations in nursing homes: a qualitative descriptive multi-site study. Age Ageing 2022 Dec 05;51(12):afac268. [doi: 10.1093/ageing/afac268] [Medline: 36580553]
42. Chez soi : le premier choix La politique de soutien à domicile. Ministère de la santé et des Services Sociaux. 2003. URL: https://publications.msss.gouv.qc.ca/msss/fichiers/2002/02-704-01.pdf [accessed 2024-08-03]
43. Documentation des soins infirmiers : Norme d'exercice. Ordre des infirmières et infirmiers du Québec. 2023. URL: https://www.oiiq.org/norme-exercice-documentation [accessed 2024-08-04]
44. Rutledge CM, O'Rourke J, Mason AM, Chike-Harris K, Behnke L, Melhado L, et al. Telehealth competencies for nursing education and practice: the four P's of telehealth. Nurse Educ 2021 Jan 19;46(5):300-305 [FREE Full text] [doi: 10.1097/NNE.0000000000000988] [Medline: 33481494]
45. Edelman LS, McConnell ES, Kennerly SM, Alderden J, Horn SD, Yap TL. Mitigating the effects of a pandemic: facilitating improved nursing home care delivery through technology. JMIR Aging 2020 May 26;3(1):e20110 [FREE Full text] [doi: 10.2196/20110] [Medline: 32412909]
46. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013 Feb 17;8(10):22-29 [FREE Full text] [doi: 10.1186/1748-5908-8-22] [Medline: 23414420]
47. Fortin MF. Foundations and stages of the research process: quantitative and qualitative methods. 2nd edition. Chenelière éducation, Montréal. URL: https://www.scirp.org/reference/referencespapers?referenceid=2631793 [accessed 2024-04-29]
48. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Mar;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
49. Sousa VL, Dourado Júnior FW, Anjos SD, Moreira AC. Nursing teleconsultation in primary health care: scoping review. Rev Lat Am Enfermagem 2024;32:e4329 [FREE Full text] [doi: 10.1590/1518-8345.7212.4329] [Medline: 39230178]
50. Hughes L, Petrella A, Phillips N, Taylor RM. Virtual care and the impact of COVID-19 on nursing: a single centre evaluation. J Adv Nurs 2022 Feb;78(2):498-509 [FREE Full text] [doi: 10.1111/jan.15050] [Medline: 34590738]
51. Mohammed HT, Hyseni L, Bui V, Gerritsen B, Fuller K, Sung J, et al. Exploring the use and challenges of implementing virtual visits during COVID-19 in primary care and lessons for sustained use. PLoS One 2021;16(6):e0253665 [FREE Full text] [doi: 10.1371/journal.pone.0253665] [Medline: 34166441]
52. Teasley SL, Sexton KA, Carroll CA, Cox KS, Riley M, Ferriell K. Improving work environment perceptions for nurses employed in a rural setting. J Rural Health 2007;23(2):179-182. [doi: 10.1111/j.1748-0361.2007.00087.x] [Medline: 17397376]
53. Li S, Jeffs L, Barwick M, Stevens B. Organizational contextual features that influence the implementation of evidence-based practices across healthcare settings: a systematic integrative review. Syst Rev 2018 May 05;7(1):72 [FREE Full text] [doi: 10.1186/s13643-018-0734-5] [Medline: 29729669]

54. Hollander JE, Carr BG. Virtually perfect? Telemedicine for COVID-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
55. Zhang T, Mosier J, Subbian V. Identifying barriers to and opportunities for telehealth implementation amidst the COVID-19 pandemic by using a human factors approach: a leap into the future of health care delivery? JMIR Hum Factors 2021 Apr 09;8(2):e24860 [FREE Full text] [doi: 10.2196/24860] [Medline: 33779566]
56. Mohammed HM, El-sol AE. Tele-nursing: opportunities for nurses to shape their profession's future. Int J Nov Res Healthc Nurs 2020;7(3):660-676 [FREE Full text]
57. Nittari G, Khuman R, Baldoni S, Pallotta G, Battineni G, Sirignano A, et al. Telemedicine practice: review of the current ethical and legal challenges. Telemed J E Health 2020 Dec;26(12):1427-1437 [FREE Full text] [doi: 10.1089/tmj.2019.0158] [Medline: 32049608]
58. Gillespie SM, Handler SM, Bardakh A. Innovation through regulation: COVID-19 and the evolving utility of telemedicine. J Am Med Dir Assoc 2020 Aug;21(8):1007-1009 [FREE Full text] [doi: 10.1016/j.jamda.2020.06.054] [Medline: 32736843]
59. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 06;8(7):e18072 [FREE Full text] [doi: 10.2196/18072] [Medline: 32442132]
60. Cannavacciuolo L, Capaldo G, Ponsiglione C. Digital innovation and organizational changes in the healthcare sector: multiple case studies of telemedicine project implementation. Technovation 2023 Feb;120:102550. [doi: 10.1016/j.technovation.2022.102550]
61. Kumar S, Tian E. Critical factors in the development, implementation and evaluation of telemedicine. In: Rashid TA, Chakraborty C, Fraser K, editors. Advances in Telemedicine for Health Monitoring: Technologies, Design and Applications. London, UK: The Institution of Engineering and Technology; 2020:1-14.
62. Beauséjour W, Hagens S. Uncovering important drivers of the increase in the use of virtual care technologies in nursing care: quantitative analysis from the 2020 national survey of Canadian nurses. JMIR Nurs 2022 Mar 31;5(1):e33586 [FREE Full text] [doi: 10.2196/33586] [Medline: 35357326]
63. Consolidated telemedicine implementation guide. World Health Organization. URL: https://www.who.int/publications/i/item/9789240059184 [accessed 2024-04-29]
64. Kho J, Gillespie N, Martin-Khan M. A systematic scoping review of change management practices used for telemedicine service implementations. BMC Health Serv Res 2020 Sep 01;20(1):815 [FREE Full text] [doi: 10.1186/s12913-020-05657-w] [Medline: 32873295]
65. Fennelly O, Cunningham C, Grogan L, Cronin H, O'Shea C, Roche M, et al. Successfully implementing a national electronic health record: a rapid umbrella review. Int J Med Inform 2020 Dec;144:104281 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104281] [Medline: 33017724]
66. Ifinedo P. Empirical study of nova scotia nurses' adoption of healthcare information systems: implications for management and policy-making. Int J Health Policy Manag 2018 Apr 01;7(4):317-327 [FREE Full text] [doi: 10.15171/ijhpm.2017.96] [Medline: 29626399]
67. Sriharan A, Sekercioglu N, Mitchell C, Senkaiahliyan S, Hertelendy A, Porter T, et al. Leadership for AI transformation in health care organization: scoping review. J Med Internet Res 2024 Aug 14;26:e54556 [FREE Full text] [doi: 10.2196/54556] [Medline: 39009038]
68. Pfitzer E, Bitomsky L, Nißen M, Kausch C, Kowatsch T. Success factors of growth-stage digital health companies: systematic literature review. J Med Internet Res 2024 Dec 11;26:e60473 [FREE Full text] [doi: 10.2196/60473] [Medline: 39661978]
69. Greenhalgh T, Ladds E, Hughes G, Moore L, Wherton J, Shaw SE, et al. Why do GPs rarely do video consultations? Qualitative study in UK general practice. Br J Gen Pract 2022 May;72(718):e351-e360 [FREE Full text] [doi: 10.3399/BJGP.2021.0658] [Medline: 35256385]
70. Lin CC, Dievler A, Robbins C, Sripipatana A, Quinn M, Nair S. Telehealth in health centers: key adoption factors, barriers, and opportunities. Health Aff (Millwood) 2018 Dec;37(12):1967-1974. [doi: 10.1377/hlthaff.2018.05125] [Medline: 30633683]
71. Garber K, Gustin T. Telehealth education: impact on provider experience and adoption. Nurse Educ 2022;47(2):75-80 [FREE Full text] [doi: 10.1097/NNE.0000000000001103] [Medline: 34628435]
72. James S, Ashley C, Williams A, Desborough J, Mcinnes S, Calma K, et al. Experiences of Australian primary healthcare nurses in using telehealth during COVID-19: a qualitative study. BMJ Open 2021 Aug 06;11(8):e049095 [FREE Full text] [doi: 10.1136/bmjopen-2021-049095] [Medline: 34362804]
73. Esmaeilzadeh P, Mirzaei T. Using electronic health records to mitigate workplace burnout among clinicians during the COVID-19 pandemic: field study in Iran. JMIR Med Inform 2021 Jun 03;9(6):e28497 [FREE Full text] [doi: 10.2196/28497] [Medline: 34033578]
74. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
75. Opoku D, Scott P, Quentin W. Healthcare professionals' perceptions of the benefits and challenges of a teleconsultation service in the Amansie-west district of Ghana. Telemed J E Health 2015 Sep;21(9):748-755. [doi: 10.1089/tmj.2014.0210] [Medline: 25942579]

76. Gagnon MP, Paré G, Pollender H, Duplantie J, Côté J, Fortin JP, et al. Supporting work practices through telehealth: impact on nurses in peripheral regions. BMC Health Serv Res 2011 Feb 04;11:27 [FREE Full text] [doi: 10.1186/1472-6963-11-27] [Medline: 21294882]
77. White M, Butterworth T, Wells JS. Reported implementation lessons from a national quality improvement initiative; productive ward: releasing time to Care™. A qualitative, ward-based team perspective. J Nurs Manag 2017 Oct;25(7):519-530. [doi: 10.1111/jonm.12489] [Medline: 28799269]
78. Rafferty AE, Jimmieson NL, Armenakis AA. Change readiness: a multilevel review. J Manag 2012 Sep 05;39(1):110-135. [doi: 10.1177/0149206312457417]
79. Meurs M, Keuper J, Sankatsing V, Batenburg R, van Tuyl L. "Get used to the fact that some of the care is really going to take place in a different way": general practitioners' experiences with e-health during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Apr 22;19(9):5120 [FREE Full text] [doi: 10.3390/ijerph19095120] [Medline: 35564519]
80. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
81. Fronczek AE, Rouhana NA, Kitchin JM. Enhancing telehealth education in nursing: applying King's conceptual framework and theory of goal attainment. Nurs Sci Q 2017 Jul 21;30(3):209-213. [doi: 10.1177/0894318417708418] [Medline: 28899265]
82. van Houwelingen CT, Moerman AH, Ettema RG, Kort HS, Ten Cate O. Competencies required for nursing telehealth activities: a Delphi-study. Nurse Educ Today 2016 Apr;39:50-62 [FREE Full text] [doi: 10.1016/j.nedt.2015.12.025] [Medline: 27006033]
83. Nilsen P, Seing I, Ericsson C, Birken SA, Schildmeijer K. Characteristics of successful changes in health care organizations: an interview study with physicians, registered nurses and assistant nurses. BMC Health Serv Res 2020 Feb 27;20(1):147 [FREE Full text] [doi: 10.1186/s12913-020-4999-8] [Medline: 32106847]
84. Arem H, Moses J, Cisneros C, Blondeau B, Nekhlyudov L, Killackey M, et al. Cancer provider and survivor experiences with telehealth during the COVID-19 pandemic. JCO Oncol Pract 2022 Apr;18(4):e452-e461. [doi: 10.1200/OP.21.00401] [Medline: 34714706]
85. Srinivasan M, Asch S, Vilendrer S, Thomas SC, Bajra R, Barman L, et al. Qualitative assessment of rapid system transformation to primary care video visits at an academic medical center. Ann Intern Med 2020 Oct 06;173(7):527-535. [doi: 10.7326/m20-1814]
86. Belcher RH, Phillips J, Virgin F, Werkhaven J, Whigham A, Wilcox L, et al. Pediatric otolaryngology telehealth in response to COVID-19 pandemic: lessons learned and impact on the future management of pediatric patients. Ann Otol Rhinol Laryngol 2021 Jul;130(7):788-795 [FREE Full text] [doi: 10.1177/0003489420976163] [Medline: 33238714]
87. Franzosa E, Gorbenko K, Brody AA, Leff B, Ritchie CS, Kinosian B, et al. "There is something very personal about seeing someone's face": provider perceptions of video visits in home-based primary care during COVID-19. J Appl Gerontol 2021 Nov;40(11):1417-1424 [FREE Full text] [doi: 10.1177/07334648211028393] [Medline: 34210200]
88. Gilkey MB, Kong WY, Huang Q, Grabert BK, Thompson P, Brewer NT. Using telehealth to deliver primary care to adolescents during and after the COVID-19 pandemic: national survey study of US primary care professionals. J Med Internet Res 2021 Sep 10;23(9):e31240 [FREE Full text] [doi: 10.2196/31240] [Medline: 34406974]
89. Connolly SL, Gifford AL, Miller CJ, Bauer MS, Lehmann LS, Charness ME. Provider perceptions of virtual care during the coronavirus disease 2019 pandemic: a multispecialty survey study. Med Care 2021 Jul 01;59(7):646-652 [FREE Full text] [doi: 10.1097/MLR.0000000000001562] [Medline: 34009880]
90. Chen W, Flanagan A, Nippak PM, Nicin M, Sinha SK. Understanding the experience of geriatric care professionals in using telemedicine to care for older patients in response to the COVID-19 pandemic: mixed methods study. JMIR Aging 2022 Aug 10;5(3):e34952 [FREE Full text] [doi: 10.2196/34952] [Medline: 35830331]
91. Pilosof NP, Barrett M, Oborn E, Barkai G, Pessach IM, Zimlichman E. Telemedicine implementation in COVID-19 ICU: balancing physical and virtual forms of visibility. HERD 2021 Jul 02;14(3):34-48 [FREE Full text] [doi: 10.1177/19375867211009225] [Medline: 34075789]
92. Kendzerska T, Zhu DT, Gershon AS, Edwards JD, Peixoto C, Robillard R, et al. The effects of the health system response to the COVID-19 pandemic on chronic disease management: a narrative review. Risk Manag Healthc Policy 2021;14:575-584 [FREE Full text] [doi: 10.2147/RMHP.S293471] [Medline: 33623448]
93. Park HY, Kwon YM, Jun HR, Jung SE, Kwon SY. Satisfaction survey of patients and medical staff for telephone-based telemedicine during hospital closing due to COVID-19 transmission. Telemed J E Health 2021 Jul;27(7):724-732 [FREE Full text] [doi: 10.1089/tmj.2020.0369] [Medline: 33216710]
94. Hunter SB, Dopp AR, Ober AJ, Uscher-Pines L. Clinician perspectives on methadone service delivery and the use of telemedicine during the COVID-19 pandemic: a qualitative study. J Subst Abuse Treat 2021 May;124:108288 [FREE Full text] [doi: 10.1016/j.jsat.2021.108288] [Medline: 33771285]
95. Lehoux P, Sicotte C, Denis JL, Berg M, Lacroix A. The theory of use behind telemedicine: how compatible with physicians' clinical routines? Soc Sci Med 2002 Mar;54(6):889-904. [doi: 10.1016/s0277-9536(01)00063-6] [Medline: 11996023]

96. Safaeinili N, Vilendrer S, Williamson E, Zhao Z, Brown-Johnson C, Asch SM, et al. Inpatient telemedicine implementation as an infection control response to COVID-19: qualitative process evaluation study. JMIR Form Res 2021 Jun 16;5(6):e26452 [FREE Full text] [doi: 10.2196/26452] [Medline: 34033576]

Abbreviations

CHSLD: centre d'hébergement et de soins de longue durée
MSSS: Ministère de la Santé et des Services sociaux
ND: nursing department
SAPA: Support Program for the Autonomy of Seniors

Edited by AF Abdulai; submitted 30.01.25; peer-reviewed by R Tripathi, T David, C Udensi; comments to author 23.03.25; revised version received 26.03.25; accepted 12.04.25; published 07.05.25.

Please cite as:
Nabelsi V, Plouffe V, Leclerc MC
Barriers to and Facilitators of Implementing Overnight Nursing Teleconsultation in Small, Rural Long-Term Care Facilities: Qualitative Interview Study
JMIR Aging 2025;8:e71950
URL: https://aging.jmir.org/2025/1/e71950
doi:10.2196/71950
PMID:

Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020)

Jingjing Chu[1], MSc; Ying Li[1], MSc; Xinyi Wang[1], BMed; Qun Xu[2], MSc; Zherong Xu[1], MD

[1]The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

[2]Zhejiang University School of Medicine, Hangzhou, China

Corresponding Author:
Zherong Xu, MD
The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

Abstract

Background: Disability profoundly affects older adults' quality of life and imposes considerable burdens on health care systems in China's aging society. Timely predictive models are essential for early intervention.

Objective: We aimed to build effective predictive models of disability for early intervention and management in older adults in China, integrating physical, cognitive, physiological, and psychological factors.

Methods: Data from the China Health and Retirement Longitudinal Study (CHARLS), spanning from 2015 to 2020 and involving 2450 older individuals initially in good health, were analyzed. The dataset was randomly divided into a training set with 70% data and a testing set with 30% data. LASSO regression with 10-fold cross-validation identified key predictors, which were then used to develop an Extreme Gradient Boosting (XGBoost) model. Model performance was evaluated using receiever operating characteristic curves, calibration curves, and clinical decision and impact curves. Variable contributions were interpreted using SHapley Additive exPlanations (SHAP) values.

Results: LASSO regression was used to evaluate 36 potential predictors, resulting in a model incorporating 9 key variables: age, hand grip strength, standing balance, the 5-repetition chair stand test (CS-5), pain, depression, cognition, respiratory function, and comorbidities. The XGBoost model demonstrated an area under the curve of 0.846 (95% CI 0.825 - 0.866) for the training set and 0.698 (95% CI 0.654 - 0.743) for the testing set. Calibration curves demonstrated reliable predictive accuracy, with mean absolute errors of 0.001 and 0.011 for the training and testing sets, respectively. Clinical decision and impact curves demonstrated significant utility across risk thresholds. SHAP analysis identified pain, respiratory function, and age as top predictors, highlighting their substantial roles in disability risk. Hand grip and the CS-5 also significantly influenced the model. A web-based application was developed for personalized risk assessment and decision-making.

Conclusion: A reliable predictive model for 5-year disability risk in Chinese older adults was developed and validated. This model enables the identification of high-risk individuals, supports early interventions, and optimizes resource allocation. Future efforts will focus on updating the model with new CHARLS data and validating it with external datasets.

(JMIR Aging 2025;8:e66723) doi:10.2196/66723

KEYWORDS

disability; prediction model; older adults; China Health and Retirement Longitudinal Study; CHARLS; medical resources allocation

Introduction

The aging of the population presents a significant global challenge, with profound implications for health care systems, economic stability, and social services [1,2]. In China, where the older population is rapidly increasing, the prevalence of disability among older adults has become a pressing concern. According to the Chinese Centers for Disease Control and Prevention, the number of older individuals with disabilities reached 52.71 million in 2020 and is projected to exceed 77.65 million by 2030. By 2030, disabled older adults are expected to account for over 57% of the total disabled population, potentially rising to more than 70% by 2050 if no preventive measures are implemented [3]. Disability in older adults encompasses various limitations or difficulties in performing daily activities independently, typically due to chronic degenerative changes in function. It is commonly assessed through activities of daily living (ADL) and instrumental activities of daily living (IADL). This growing burden of disability impacts quality of life and places a significant strain on families and public resources. Therefore, accurate prediction of disability is crucial for early intervention and effective management.

Despite significant research efforts to forecast disability in older adults, existing models often lack sufficient precision and fail to account for the complex, multifactorial nature of disability.

These models typically overlook the broader context of risk factors and offer limited utility for public health decision-making [4-6]. For instance, a study by Sun et al [5] identified depressive symptoms as a significant predictor across different types of disability. However, many existing models still fail to incorporate mental health factors alongside physical health indicators, limiting their real-world applicability. Furthermore, cognitive impairment has been found to be a strong predictor of disability in specific ADL and IADL tasks [6], highlighting the importance of integrating mental, cognitive, and physical health factors in predictive models. Disability, as a complex health issue, is influenced by multiple risk domains, including chronic diseases, polypharmacy, aging, mental health problems, unhealthy lifestyles, and the family social environment. A comprehensive predictive model should serve as a vital tool to improve early identification of at-risk individuals, inform public health strategies, and optimize resource allocation.

This study seeks to address the limitations in existing research by developing a disability prediction model specifically designed for the older Chinese population, using longitudinal data from the China Health and Retirement Longitudinal Study (CHARLS) collected between 2015 and 2020 [7]. While previous research has developed disability prediction models based on CHARLS data from 2015 to 2018 [8], the release of the 2020 survey data enables the extension of the analysis over a longer time frame. This study will leverage the validated Extreme Gradient Boosting (XGBoost) algorithm [8] to explore disability predictors in greater depth. In addition, by integrating variables such as sarcopenia and frailty-related indicators, which have previously been underexplored in predictive models for disability, we offer a more nuanced understanding of the physical, cognitive, physiological, and psychological factors contributing to disability risk. We aim to create a predictive model that not only offers high precision but also provides practical insights for health care professionals and policymakers.

Methods

Study Population

The data for this study were sourced from the CHARLS, initiated in 2011 by the National School of Development at Peking University. The CHARLS used a stratified, multistage Probability Proportional to Size random sampling method, covering 150 counties and 450 villages and urban communities across 28 provinces, involving 17,708 individuals from 10,257 households. Follow-up surveys were conducted in 2013, 2015, 2018, and 2020, with detailed methodology available in other publications [7].

Initially, 21,095 participants from the 2015 baseline survey were included. The final cohort consisted of 2450 individuals after applying the following exclusion criteria: (1) no information on biomarker or blood data; (2) younger than 60 years (for this study, older adults were defined as individuals aged 60 years or older, in accordance with the World Health Organization and the Chinese government’s standard for aging population classification); (3) missing ADL or IADL data in 2015; (4) missing follow-up ADL or IADL data in 2020; (5) having ADL and IADL limitations or any form of disability in 2015, including physical, intellectual, visual, auditory, or significant speech impairments; and (6) other relevant data missing.

Ethical Considerations

The Institutional Review Board of Peking University (IRB No. IRB00001052-11014) approved the research, and all respondents provided informed consent. CHARLS adheres to the Declaration of Helsinki and China's Personal Information Protection Law. The CHARLS database adheres to strict privacy protection and anonymization principles during data collection and processing to ensure the security of participants' personal information.

Assessment of Disability

Disability in this study was defined as impairment in performing ADL and IADL, which is commonly used in geriatric research to evaluate functional limitations in older adults. ADL assessed the ability to perform fundamental self-care tasks such as dressing, bathing, eating, getting out of bed, toileting, and managing urination and bowel movements. IADL measured more complex daily tasks, including household chores, cooking, shopping, phone use, financial management, and medication adherence. Responses were categorized into four levels: (1) no difficulty, (2) difficulty but can still do it, (3) difficulty and need help, and (4) cannot do it. To create a binary outcome, responses of (2) “Difficulty but can still do it,” (3) “Difficulty and need help,” and (4) “Cannot do it” were coded as 1 (indicating ADL and IADL disability), while the response “No difficulty” was coded as 0 (no disability). Participants were classified as having ADL and IADL disabilities if they reported any level of difficulty (levels 2 - 4) in at least one ADL or IADL item [9].

Predictive Variables

Clinical Factors

Laboratory assessments included a range of biomarkers: white blood cell count, hemoglobin, hematocrit, triglycerides, total cholesterol, glucose, uric acid, creatinine, blood urea nitrogen (BUN), high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, cystatin C, C-reactive protein, and glycated hemoglobin.

Depressive symptoms were evaluated using the 10-item Center for Epidemiologic Studies Depression Scale (CES-D) [10]. CES-D scores ranged from 0 to 30, with higher scores indicating more severe symptoms. The CES-D has been validated for Chinese middle-aged and older populations [11]. A score of 10 or above was used to indicate depression, while scores below 10 indicated no depressive symptoms [12]. Cognitive function was assessed using a modified version of the telephone interview for cognitive status (TICS) questionnaire [13]. The overall cognitive score was calculated by summing the scores from four domains: (1) orientation (5 points), (2) computation (5 points), (3) memory (20 points), and (4) drawing (1 point), with a total possible score of 31 points [14]. Higher scores indicate better cognitive performance.

Physical Performance

The physical examination included measurements of systolic and diastolic blood pressure, pulse, and respiratory function. Respiratory function was measured using a peak flow meter.

Participants were instructed to stand, take a deep breath, and blow as hard and fast as possible into the mouthpiece. The highest value from 3 attempts was recorded for analysis. Physical performance was assessed using gait speed, the 5-repetition chair stand test (CS-5), and standing balance. Gait speed was measured to evaluate lower limb function and mobility. A walking course of 2.5 meters was set up, and participants were instructed to walk the course twice at their usual pace. The average gait speed was calculated by dividing the distance by the time taken. The 5-repetition CS-5 was conducted to assess lower limb strength and endurance. Participants were asked to sit in a chair with their arms folded across their chest and, upon the examiner's command, to stand up and sit down 5 times consecutively at their fastest pace without using their arms for support. The total time required to complete the 5 repetitions was recorded, with a longer duration indicating poorer lower limb function. The standing balance assessment involved maintaining a standing position for 10 seconds in three distinct foot placements: (1) side-by-side, (2) semitandem, and (3) full tandem. Handgrip strength, the primary indicator of muscle strength, was measured for each participant using a Yuejian WL-1000 dynamometer (Nantong Yuejian Physical Measurement Instrument Co). Handgrip strength was measured in both the dominant and nondominant hands, with 2 measurements per hand. The higher value for each hand was recorded, and the average value for the 2 hands was taken to represent the handgrip strength. Together, these assessments provided a comprehensive evaluation of physical function and performance.

Appendicular skeletal muscle mass (ASM) was estimated using a formula specifically developed for the Chinese population, which closely corresponds with dual-energy X-ray absorptiometry measurements [15,16]. The formula accounts for weight, height, sex (1 for males, 2 for females), and age as follows:

$$ASM = 0.193 \times \text{weight (kg)} + 0.107 \times \text{height (cm)} - 4.157 \times \text{sex} - 0.037 \times \text{age} - 2.631$$

Potential Covariates

Covariates for our study were identified from previous literature and grouped into 2 main categories. The first included social and lifestyle factors: age, gender, BMI, marital status, residential area, daily sleep hours, and alcohol and tobacco use. BMI categories were defined as underweight (BMI <18.5 kg/m^2), normal weight (BMI =18.5 - 24 kg/m^2), and overweight (BMI ≥24 kg/m^2). The second category addressed pain, incidents of falling, and number of comorbidities (hypertension, dyslipidemia, diabetes, cancer, stroke, heart disease, lung disease, liver disease, kidney disease, digestive disease, mental health disorders, memory disorders, asthma, and arthritis) [14,17]. The comorbidity classification was based on the CHARLS questionnaire design. Neurological disorders, including Parkinson disease and Alzheimer disease, are included under memory-related diseases.

Statistical Analysis

The preprocessed dataset was split into a training subset with 70% data and a testing subset with 30% data. Continuous variables were described using medians and IQR, with comparisons using the Mann–Whitney *U* test. Count variables were expressed as frequencies and percentages and assessed using the χ^2 test. Model development and testing were performed using the training and testing sets, respectively.

Initial correlation analysis identified potential multicollinearity. Variable selection was conducted exclusively on the training set using LASSO regression with 10-fold cross-validation to prevent information leakage and ensure an unbiased evaluation of model performance. LASSO regression was chosen over other methods due to its ability to perform simultaneous variable selection and regularization, reduce overfitting, and enhance model interpretability. In addition, ablation experiments were conducted to evaluate the effect of removing specific features related to sarcopenia and frailty on model performance. The logloss metric was used to assess the performance of the models with and without these features. The selected variables informed the development of an XGBoost model, a machine learning algorithm that uses gradient boosting through decision trees to iteratively minimize prediction errors. For the optimization of the XGBoost model, hyperparameter tuning was performed using a grid search approach. Key hyperparameters were tuned, including the number of boosting rounds (nrounds), which determines the number of iterations for boosting, and the maximum tree depth, which controls the complexity of each individual tree. The learning rate (eta) was adjusted to control the weight of each update during training. In addition, the minimum loss reduction (gamma) for tree splitting, the feature subsampling ratio (colsample_bytree), and the minimum child weight were optimized to control the model's complexity and prevent overfitting. The subsample ratio (subsample) was also tuned to control the fraction of training data used in each boosting round. The optimal parameters were selected based on the lowest logloss value obtained during cross-validation.

Model performance was assessed using receiver operating characteristic (ROC) curves and area under the curve (AUC) values, with higher AUC indicating better discrimination. Calibration curves evaluated the agreement between predicted and observed outcomes. Decision curve analysis (DCA) and clinical impact curves (CIC) aided in determining optimal application and estimating the model's impact on patient management. SHapley Additive exPlanations (SHAP) values were used to interpret variable importance and model transparency. Four key SHAP plots were generated: (1) a summary plot, (2) dependence plot, (3) interaction plot, and (4) force plot. The model has been deployed on a web-based platform.

Analyses were conducted with R software (R Foundation for Statistical Computing), version 4.3.2. A *P* value <.05 was considered statistically significant.

Results

Baseline Characteristics

This study assessed a cohort of 2450 older adults initially in good health. Over a 5-year follow-up period, 610 participants developed disabilities, resulting in a disability incidence rate of

24.90%. The dataset was split 7:3, with the training set consisting of 1715 individuals (427 with disabilities) and the testing set comprising 735 participants (183 with disabilities). Baseline characteristics of both sets were detailed in Table 1 and Figure 1. Except for differences in sleep duration and white blood cell counts, no statistically significant differences were observed between the 2 groups ($P>.05$).

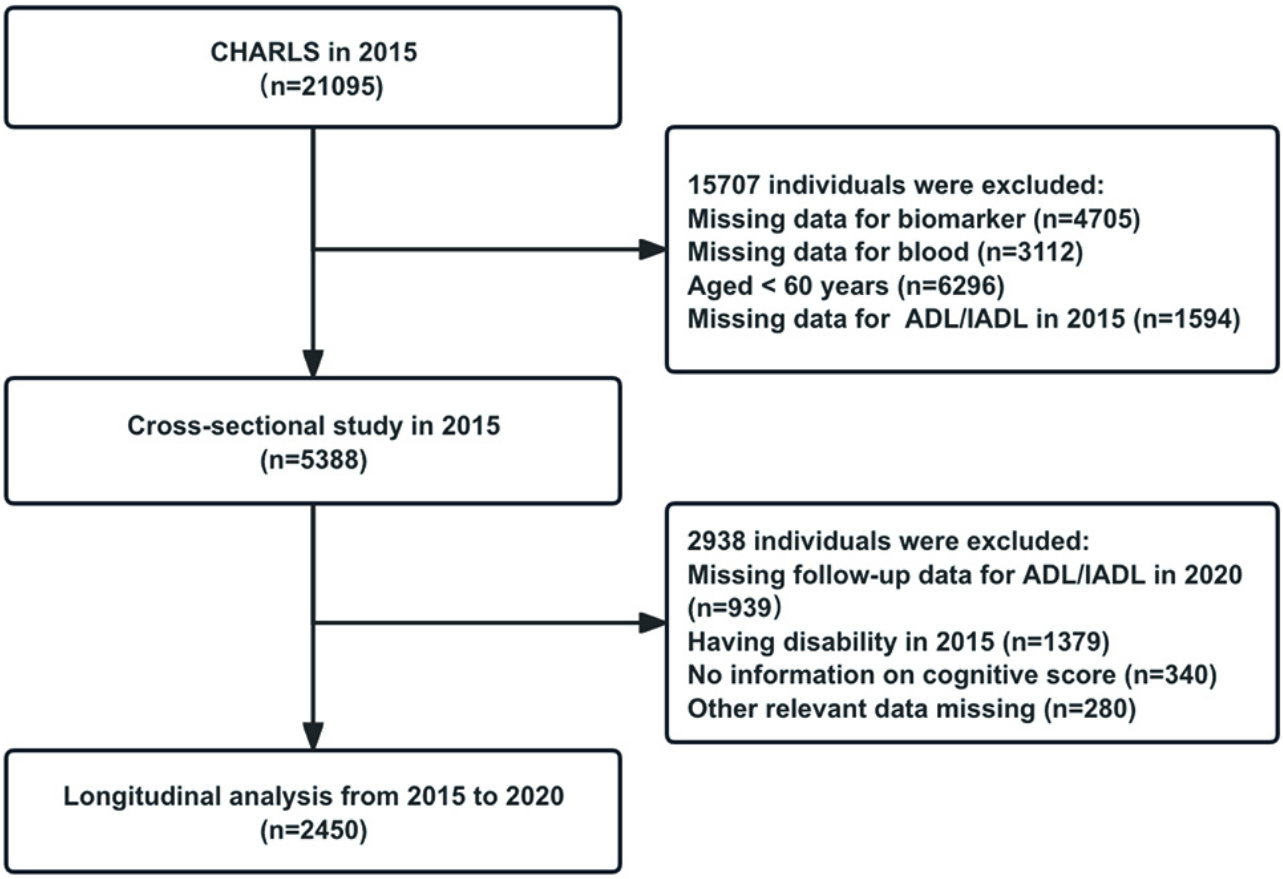
Table . Baseline characteristics in the training and testing subset.

Variable	Total	Training set	Testing set	*P* value
	N=2450	N=1715	N=735	
Dependent variable, n (%)				
Disability	610 (24.90)	427 (24.90)	183 (24.90)	—
Non-disability	1840 (75.10)	1288 (75.10)	552 (75.10)	—
ASM/Ht^2	6.89 [5.97-7.59]	6.87 [5.96-7.56]	6.92 [6-7.70]	.12
Hand grip (kg)	31 [25-38]	31 [25-38]	31 [24.60-38]	.78
Gait speed (m/s)	0.83 [0.69-0.97]	0.83 [0.69-0.97]	0.83 [0.70-0.99]	.17
CS-5 (s)	8.78 [7.22-10.60]	8.81 [7.30-10.70]	8.71 [7.08-10.50]	.11
Balance	3 [3-3]	3 [3-3]	3 [3-3]	.99
Age (year)	65 [62-70]	65 [62-70]	65 [62-70]	.70
Sex, n (%)				.50
Male	1393 (56.86)	967 (56.38)	426 (57.96)	
Female	1057 (43.14)	748 (43.62)	309 (42.04)	
BMI level, n (%)				.14
Underweight	142 (5.80)	104 (6.06)	38 (5.17)	
Normal weight	1299 (53.02)	926 (54.00)	373 (50.75)	
Overweight	1009 (41.18)	685 (39.94)	324 (44.08)	
Marital status, n (%)				.21
Unmarried	417 (17.02)	303 (17.67)	114 (15.51)	
Married	2033 (82.98)	1412 (82.33)	621 (84.49)	
Living area, n (%)				.59
Rural	1958 (79.92)	1376 (80.23)	582 (79.18)	
Urban	492 (20.08)	339 (19.77)	153 (20.82)	
Education level, n (%)				.24
Illiterate	811 (33.10)	577 (33.64)	234 (31.84)	
Primary school	1062 (43.35)	753 (43.91)	309 (42.04)	
Middle school	388 (15.84)	256 (14.93)	132 (17.96)	
High school and above	189 (7.71)	129 (7.52)	60 (8.16)	
Sleeping, n (%)				.04[a]
<6 h	699 (28.53)	511 (29.80)	188 (25.58)	
≥6 h	1751 (71.47)	1204 (70.20)	547 (74.42)	
Pain, n (%):				.17
No	1981 (80.86)	1374 (80.12)	607 (82.59)	
Yes	469 (19.14)	341 (19.88)	128 (17.41)	
Falldown, n (%)				.81
No	2102 (85.80)	1469 (85.66)	633 (86.12)	
Yes	348 (14.20)	246 (14.34)	102 (13.88)	
Smoking, n (%)				.35
No	1497 (61.10)	1037 (60.47)	460 (62.59)	
Yes	953 (38.90)	678 (39.53)	275 (37.41)	
Drinking, n (%)				.45

Variable	Total	Training set	Testing set	*P* value
	N=2450	N=1715	N=735	
No	1496 (61.06)	1056 (61.57)	440 (59.86)	
Yes	954 (38.94)	659 (38.43)	295 (40.14)	
Comorbidities, n (%)				.33
0	1705 (69.59)	1209 (70.50)	496 (67.48)	
1	538 (22.96)	365 (21.28)	173 (23.54)	
≥2	207 (8.45)	141 (8.22)	66 (8.98)	
Depression, n (%)				.37
No	1900 (77.55)	1339 (78.08)	561 (76.33)	
Yes	550 (22.45)	376 (21.92)	174 (23.67)	
Cognition	15.00 [12.00-18.00]	15.00 [11.50-18]	16 [12-19]	.46
Systolic BP[b]	129.67 [116.67-143.67]	129.67 [116.33-143.33]	130.33 [118-143.67]	.21
Diastolic BP	74.00 [67.33-81.33]	74 [67.33-81.33]	74.33 [67.33-81.83]	.71
Pulse	72.3 [66-79.3]	72.67 [66-79.67]	72 [66-79]	.53
Respiratory function	280 [203.33-360]	276.67[200-356.67]	280 [213-363.00]	.15
WBC[c] (1000)	5.70 [4.75-6.80]	5.67 [4.72-6.78]	5.90 [4.80-6.90]	.01[a]
HGB[d] (g/dl)	13.74 [12.66-14.80]	13.70 [12.60-14.80]	13.90 [12.70-14.90]	.13
HCT[e] (%)	41.60 [38.50-45.00]	41.40 [38.40-44.80]	41.90 [38.70-45.20]	.19
TG[f] (mg/dl)	110.62 [81.42-161.95]	109.73 [80.53-161.06]	114.16 [81.42-163.72]	.30
CHO[g] (mg/dl)	182.63 [161-205.79]	183.40 [161.39-206.56]	181.47 [159.85-203.86]	.24
GLU[h] (mg/dl)	97.30 [90.09-108.11]	95.50 [90.09-108.11]	97.30 [90.10-108.11]	.46
UA[i] (mg/dl)	5.00 [4.10-5.90]	5.00 [4.10-5.90]	4.90 [4.10-5.90]	.51
CRP[j] (mg/l)	1.40 [0.80-2.60]	1.40 [0.80-2.70]	1.40 [0.80-2.40]	.17
HbA1c[k] (%)	5.90 [5.60-6.20]	5.90 [5.60-6.20]	5.90 [5.60-6.20]	.81
CREA[l] (mg/dl)	0.81 [0.69-0.93]	0.80 [0.69-0.93]	0.81 [0.70-0.94]	.27
BUN[m] (mg/dl)	15.13 [12.89-18.49]	15.41 [12.89-18.49]	15.13 [12.61-18.49]	.37
HDL[n] (mg/dl)	50.19 [43.24-57.92]	50.19 [43.24-58.30]	50.19 [43.24-57.14]	.50
LDL[o] (mg/dl)	101.93 [84.56-120.85]	102.32 [84.56-121.24]	99.61 [84.17-120.08]	.15

[a]*P*< .05
[b]BP: blood pressure.
[c]WBC: white blood cell.
[d]HGB: hemoglobin.
[e]HCT: hematocrit.
[f]TG: triglycerides.
[g]CHO: total cholesterol.
[h]GLU: glucose.
[i]UA: uric acid.
[j]CRP: C-reactive protein.
[k]HbA1c: glycated hemoglobin.
[l]CREA: creatinine.
[m]BUN: blood urea nitrogen.
[n]HDL: high-density lipoprotein cholesterol.
[o]LDL: low-density lipoprotein cholesterol.

Figure 1. Flowchart of the data selection. CHARLS: China health and retirement longitudinal study, ADL/IADL: Activities of daily living and instrumental activities of daily living.



Predictor Selection

Figure 2 represented the interrelations among the continuous independent variables measured in the study. The matrix used varying shades of color and circle sizes to illustrate the magnitude and direction of correlation coefficients. The analysis revealed significant correlations, such as a negative association between hand grip and age, and a positive association between CHO and high-density lipoprotein (P<.001).

To identify the strongest predictors of disabilities, the training dataset was normalized to account for different measurement units across variables. With disability as the dependent variable, 36 potential predictors were evaluated using LASSO regression. The compressive variable coefficient was used to avoid overfitting and improve predictive accuracy. The parameter λ was selected based on the largest λ within 1 SD of the minimal binomial deviance to enforce stricter penalty constraints. The LASSO regression retained 9 predictors with non-zero coefficients (Figure 3): age, hand grip, standing balance, CS-5, pain, depression, cognition, respiratory function, and the count of comorbidities.

Following LASSO regression, Table 2 summarizes the results of the ablation experiments, which evaluates the impact of removing specific sarcopenia- and frailty-related features (hand grip, CS-5, and standing balance) on model performance. Logloss was used as the primary metric to evaluate the model's performance, with lower values indicating better predictive accuracy. The results showed that removing these features increased the logloss, with the most significant increase observed when all 3 features were removed simultaneously. These findings suggest that including sarcopenia-related parameters is crucial for maintaining the model's predictive accuracy.

Figure 2. Correlation matrix of continuous independent variables. *: *P*<.05; **: *P*<.01; ***: *P*<.001. ASM/Ht2: appendicular skeletal muscle mass and height 2, CS-5: five-repetition chair stand test, Systolic and Diastolic BP: systolic blood pressure, WBC: white blood cell, HGB: hemoglobin, HCT: hematocrit, TG: triglycerides, CHO: total cholesterol, GLU: glucose, UA: uric acid, CRP: C-reactive protein, HbA1c: glycated hemoglobin, CREA: creatinine, BUN: blood urea nitrogen, HDL: high-density lipoprotein cholesterol, LDL: low-density lipoprotein cholesterol. Positive correlations are represented by blue tones, and negative correlations by red tones, with the intensity of the color indicating the strength of the correlation. Circle size is proportional to the absolute value of the correlation coefficient.

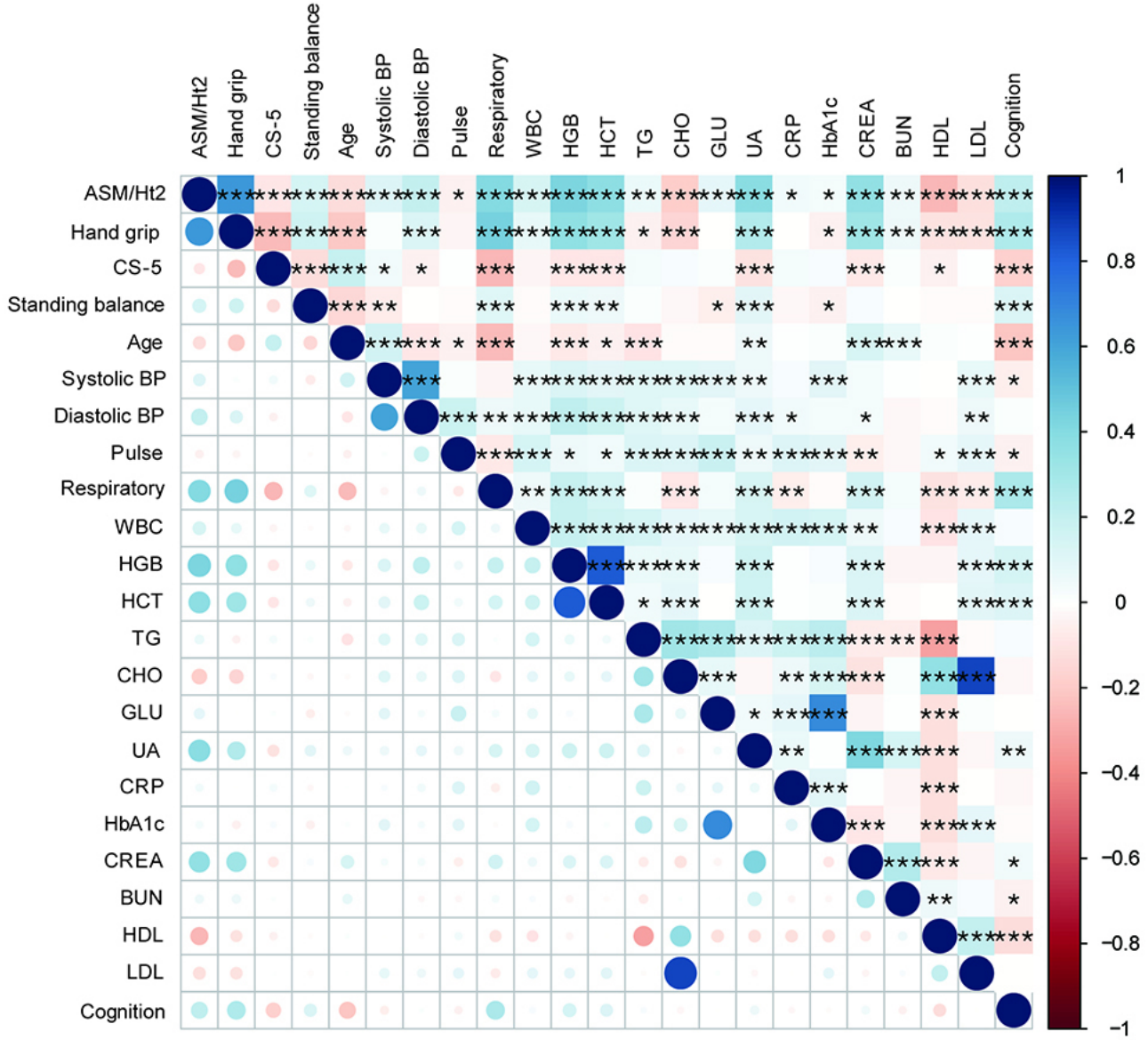


Figure 3. Variable selection via LASSO regression model. (A)Optimal parameter selection in LASSO regression. This plot illustrates the choice of the optimal λ, displaying log(λ) on the horizontal axis and regression coefficients on the vertical axis. (B)LASSO regression parameter (λ) selection via binomial deviance plot. Each point represents the model's deviance at varying log (λ) values, with the vertical dotted line indicating the λ value that minimizes the binomial deviance.

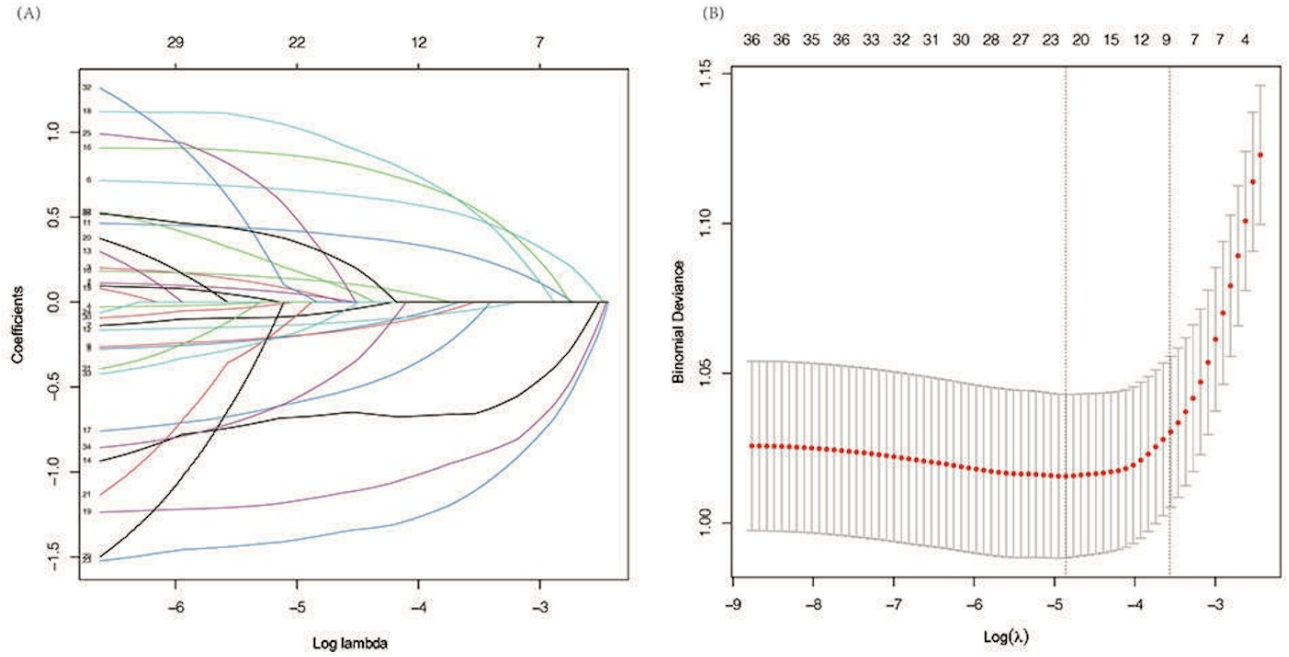


Table . Performance comparison of ablation experiments on model performance.

Model configuration	Logloss	Change	Number of features
All features (full model)	0.411	0	9
Removing hand grip	0.427	0.016	8
Removing CS-5[a]	0.421	0.010	8
Removing standing balance	0.413	0.002	8
Removing hand grip, CS-5, and standing balance	0.442	0.031	6

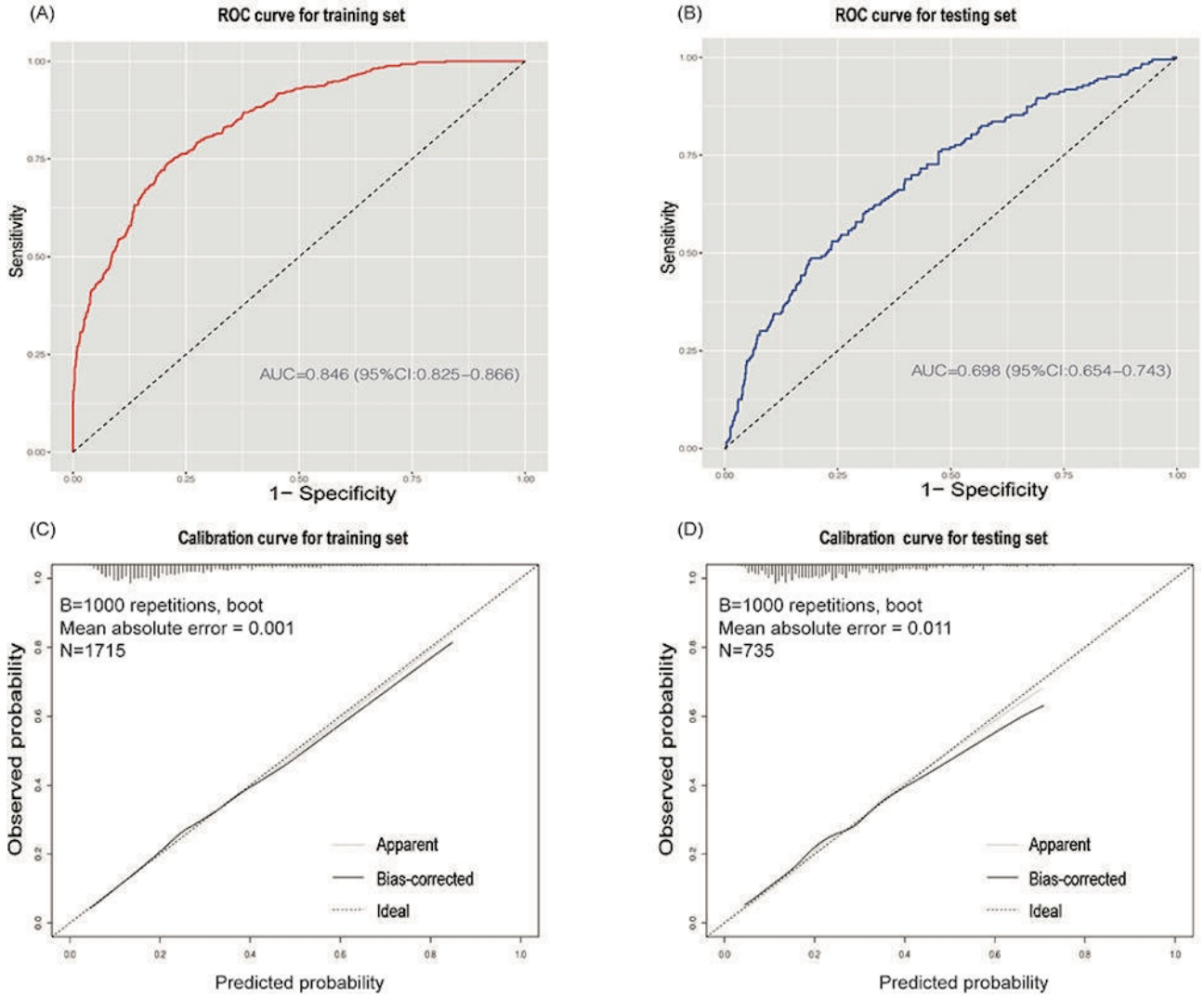
[a]CS-5: five-repetition chair stand test

Construction and Assessment of the Predictive Model

Using disability outcomes from 2020 as the dependent variable and 9 predictors selected through LASSO regression, a predictive model was constructed using the XGBoost algorithm. The model's hyperparameters were optimized through grid search and cross-validation. The best parameters identified were: nrounds=100, max_depth=3, eta=0.1, gamma=0.1, colsample_bytree=0.8, min_child_weight=3, and subsample=0.8. These parameters were used to train the final XGBoost model, which was then evaluated on the testing dataset.

The performance of the XGBoost model was evaluated using ROC curves to assess its discrimination ability. In the training set, the model achieved an AUC of 0.846 (95% CI 0.825 - 0.866), indicating good discrimination (Figure 4A). In the testing set, the AUC was 0.698 (95% CI 0.654 - 0.743), reflecting moderate predictive accuracy (Figure 4B).

Figure 4. Receiver operating characteristic curves.

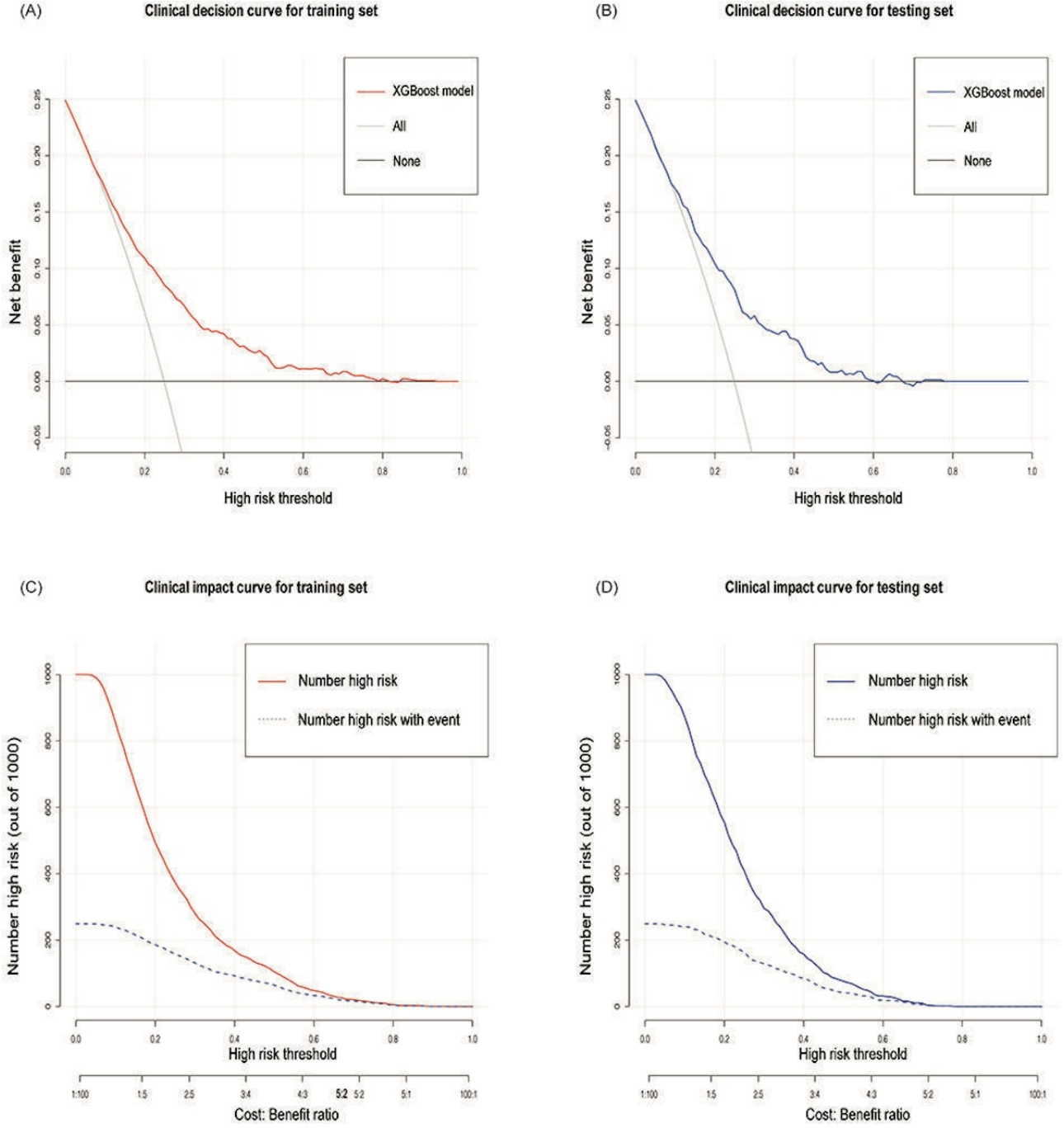


Calibration curves for the training and testing sets assessed the model's predictive accuracy. The training set showed a mean absolute error of 0.001 (Figure 4C), suggesting high precision, while the testing set had a mean absolute error of 0.011 (Figure 4D). Both curves closely approximated the ideal line, confirming the model's reliable prediction of disability risk.

Clinical decision analysis demonstrated the effectiveness of the XGBoost model in predicting disability across different risk thresholds. In the training set, the DCA showed that using the XGBoost model to identify high-risk patients provided a net benefit (Figure 5A). For example, at a chosen risk threshold of 0.30, applying the model's predictions would result in a better net benefit than treating all patients or treating none, highlighting the model's clinical utility in improving decision-making. The testing set also demonstrated a similar net benefit (Figure 5B), confirming the model's robustness and clinical applicability in an independent dataset.

In both datasets, the "Number high risk" line decreased steeply with increasing thresholds, indicating fewer individuals were classified as high-risk under stricter criteria (Figure 5C and D). In contrast, the "Number high risk with event" line, representing individuals who experienced disability, showed a more gradual decline. These trends highlight the model's ability to focus predictions on a targeted group as thresholds increase, demonstrating its utility in guiding clinical decision-making and optimizing interventions for those most likely to benefit. For a more comprehensive evaluation of the model's performance, the specificity, accuracy, positive predictive value (PPV), and negative predictive value at thresholds of 0.2 and 0.5 are provided in Multimedia Appendix 1.

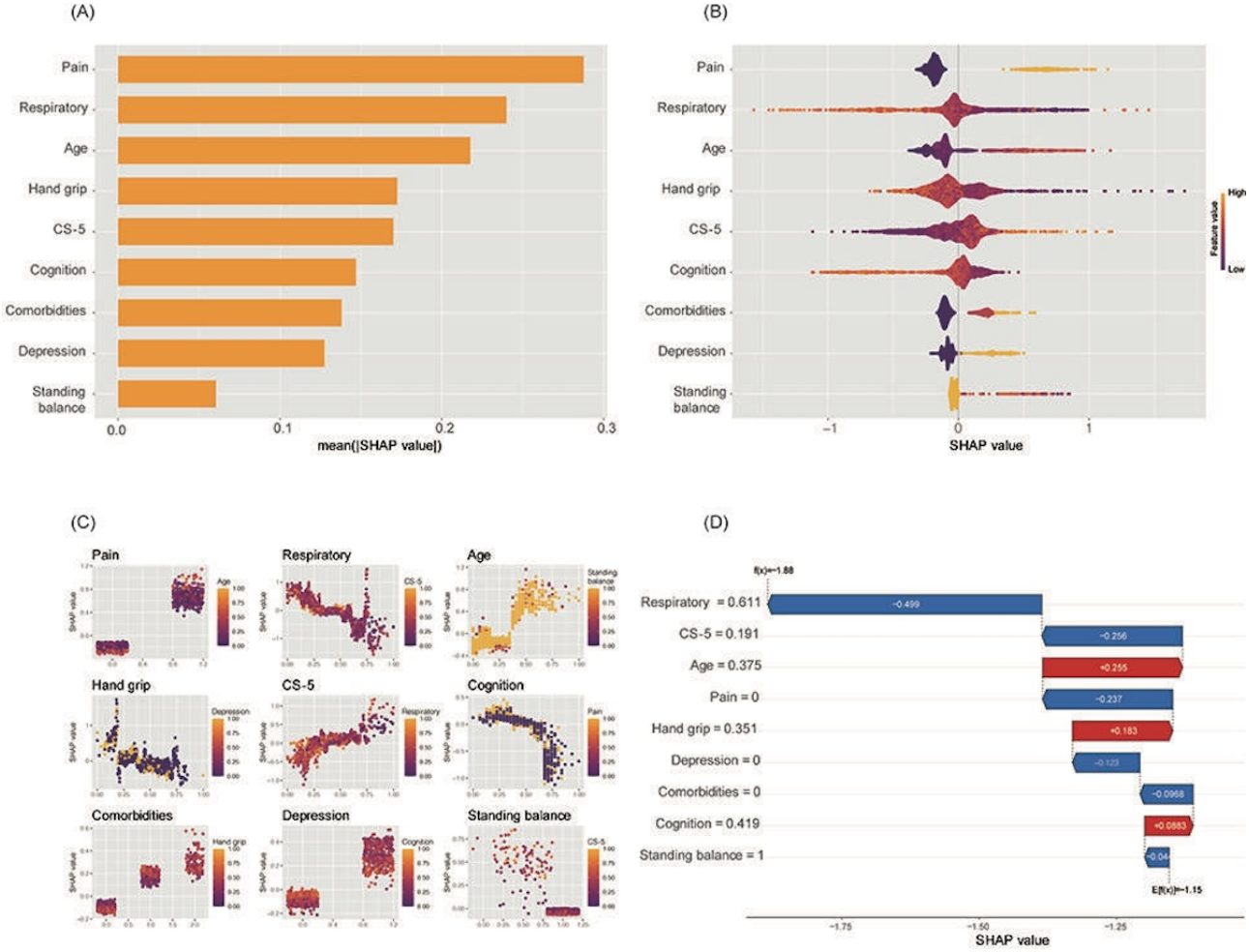
Figure 5. Clinical decision curves and impact curves for XGBoost model. (A)The red line represents the net benefit of the training set. (B)The blue line represents the net benefit of the testing set. The 'all' line indicates the benefit when all patients are treated, and the 'none' line when no patients are treated. (C-D) The solid lines depict the total number of individuals identified as high risk, and the dashed lines represent those at high risk who experienced the true event.



SHAP for Model Interpretation

We used SHAP values to assess the influence of each variable on the 5-year disability risk. Figure 6A ranked predictors by their mean SHAP values, reflecting their average contribution to the model's output. Pain had the highest mean SHAP value, followed by respiratory function and age, indicating their strong overall influence on disability prediction. Figure 6B showed the SHAP value distributions, where pain, respiratory function, and age exhibited the broadest ranges, suggesting their dynamic and individualized impact. Hand grip and CS-5 also significantly influenced the model, highlighting their importance in predicting physical function-related disability, while cognition, comorbidities, and depression showed more consistent contributions. Balance had the least impact.

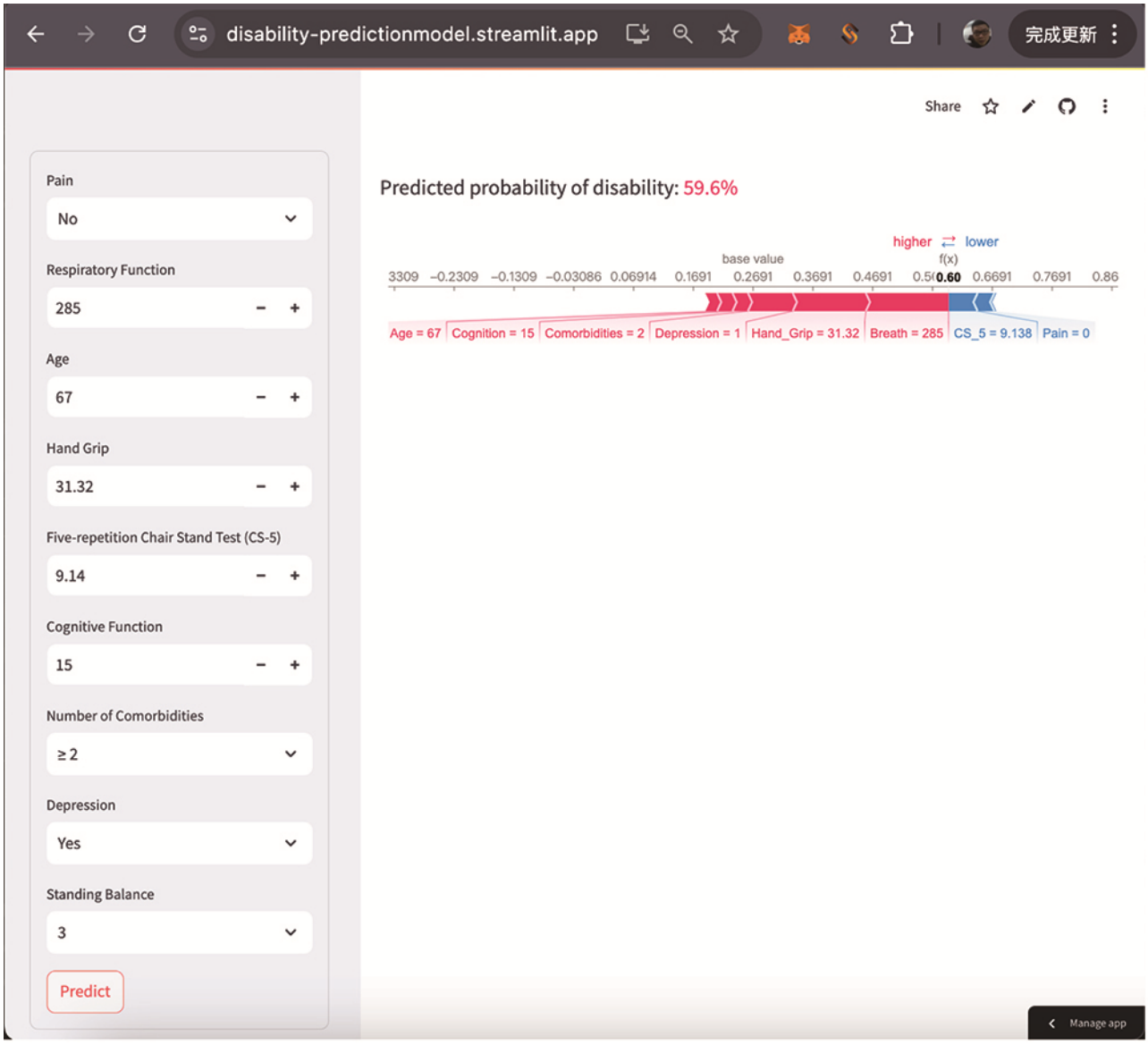
Figure 6. SHAP: SHapley additive exPlanations value interpretation diagram for predicting disability.(A)Variable importance in the predictive model as measured by SHAP values.(B)SHAP value distribution for predictive model variables. SHAP values for model variables are shown as violins. The color represents feature value intensity, and the width indicates impact density. (C)Scatterplot matrix of SHAP values for model predictors. Each plot reveals the influence of a single variable on the model output, with color intensity indicating the magnitude of the feature value. (D)Individual sample SHAP value analysis for disability risk prediction.



The SHAP summary plot (Figure 6C) provided an overview of the overall influence of each predictor on the model's output, revealing that pain, age, and respiratory function had the most substantial and wide-ranging influence on the predicted disability risk. The SHAP dependence plot (Figure 6D) visualized the individualized impact of these predictors on a single patient's disability risk profile, offering insights into the model's decision-making process at both macro and micro levels. Together, these plots provided a comprehensive understanding of the predictors' contributions to disability risk prediction.

A web-based calculator [18] enables clinicians to estimate the 5-year disability probability by entering patient-specific data [18], aiding in personalized clinical decisions (Figure 7).

Figure 7. Online calculator for individual disability prediction.



Discussion

Principal Findings and Comparison With Previous Works

This study constructed an effective 5-year disability prediction model using baseline data from CHARLS 2015 to forecast disability occurrence in 2020. The model identified 9 key predictive variables (pain, respiratory function, age, handgrip strength, CS-5, cognitive function, depression, comorbidities, and standing balance) that are closely associated with disability incidence in Chinese older adults. Calibration curves demonstrated the model's strong discrimination and consistency in both training and test sets, while DCA and CIC highlighted its positive clinical and social application value. The use of longitudinal data from CHARLS allowed for a more accurate, data-driven understanding of aging-related disability trends, leveraging demographic and health-related variables highly relevant to the Chinese context. In addition, the inclusion of sarcopenia and frailty-related diagnostic indicators as predictive variables represents a novel aspect of this research. These indicators enhance the model's sensitivity to physical, cognitive, physiological, and psychological changes associated with aging that contribute to disability. This model offers a practical tool for improving disability prevention and management in older adults.

One notable observation in this study is the difference in AUC values between the training set (0.846) and the testing set (0.698). This discrepancy suggests potential overfitting, where the model may have captured noise or random fluctuations in the training data that do not generalize to unseen data. The observed discrepancy may be due to the imbalance in the dataset with respect to the outcome variable (disability status). To address this, we attempted oversampling to balance the data. However, this approach increased model complexity by retaining additional statistically significant predictors, raising concerns about overfitting. As a result, we proceeded with the original dataset to prioritize model simplicity and reduce overfitting risks. Although the model's discriminative performance decreased in the test set, the overall trends remain robust, providing valuable insights into disability prediction for older populations. Moving forward, we plan to conduct external

validation using additional datasets or longitudinal CHARLS follow-up data to further assess the model's generalizability across diverse settings and time points. We selected the XGBoost algorithm for modeling based on its demonstrated ability to handle complex datasets and address class imbalance by assigning higher weights to the minority class [8]. This approach enhances the model's ability to predict the minority class accurately. XGBoost's advantages over other algorithms include its superior performance with structured and unstructured data, its regularization techniques to reduce overfitting, and its efficiency in training large datasets. In addition, its gradient boosting framework captures complex variable interactions, making it more suitable than traditional linear models. To optimize performance, we used advanced techniques such as cross-validation and hyperparameter tuning, including a comprehensive grid search over key parameters. This rigorous process ensured the model was robust and well-calibrated, improving its predictive performance. By extending the analysis to the latest 2020 CHARLS data, this study offers a more comprehensive 5-year prediction window compared with previous work focused on shorter time frames.

SHAP summary charts clarified the role and importance of each variable in predicting disability, providing transparency and interpretability to the model. According to the SHAP chart, pain, respiratory function, and age were the top 3 factors in importance, with the wide distribution of SHAP values. This indicates that changes in these variables significantly alter the risk of disability. Chronic pain, particularly lower back and neck pain, is a leading cause of disability globally, as highlighted by the Global Burden of Disease Study 2015. These types of pain are major causes of years lived with disability in many regions, including Latin America, the Caribbean, most regions of Asia, Oceania, and sub-Saharan Africa [19]. Chronic pain is closely linked to functional disability and poor physical performance in the older, as supported by various studies [20,21]. Respiratory function declines with age in older people, and respiratory impairment accounts for 20.7% of all types of disability [22]. Maximal inspiratory pressure and maximal expiratory pressure are correlated with hand-grip strength and skeletal muscle mass index [23,24]. Respiratory sarcopenia, characterized by a decrease in respiratory muscle strength alongside systemic skeletal muscle with aging [25], can lead to deterioration in respiratory force generation, adversely affecting activities of daily living [26]. Overall, respiratory impairment is prevalent among older individuals and is linked to physical inactivity and poor performance-based mobility [27]. Age is an independent risk factor for disability, with intrinsic capacity and functional ability declining with age. Disability levels are highest in the oldest patients [28], and age correlates with increased pain and respiratory impairment. Older adults are more likely to experience these issues, further increasing their disability risk [29]. In summary, higher pain scores, poorer respiratory function, and older age are associated with a greater risk of disability. Clinically, this suggests the need for emphasis on pain management and respiratory exercises in old people, particularly for those with chronic respiratory diseases.

The concentrated distribution of SHAP values for handgrip strength, CS-5, and cognitive function indicates these variables significantly influenced disability prediction. The ablation experiments further confirmed the impact of handgrip strength and CS-5 on the model. When these sarcopenia-related features were removed, the model's performance was notably affected, as indicated by a significant increase in logloss. This finding aligns with the SHAP analysis results, which showed a wide distribution of values for these variables. Weak handgrip strength is identified as a key component of sarcopenia, strongly associated with subsequent disability and mortality [30]. Reduced handgrip strength and lower extremity strength, as measured by the CS-5, are strong predictors of functional impairment, disability, and low health-related quality of life, significantly increasing the risk of severe disability, frailty, and other health limitations in older adults [31,32]. Cognitive function is another crucial risk factor for disability. Studies have shown that cognitive decline is associated with ADL disability [33], and longitudinal research indicates that cognitive impairment may precede ADL disability, serving as a predictor of intermediate and late-stage ADL loss [34]. Physical and cognitive functions are closely related, with physical activity enhancing neurogenesis in the adult brain. Dual-task training, which enhances both cognitive and physical functions, has shown positive effects on cognitive function and physical activity in older individuals [35]. In summary, declines in handgrip strength, lower extremity strength (CS-5), and cognitive function are positively correlated with an increased risk of disability. This underscores the importance of targeted interventions, such as early muscle strength training for the upper and lower limbs and cognitive function exercises, to help reduce the risk of disability.

Although the roles of depression, comorbidities, and standing balance are less significant in predicting disability, they still contribute to its progression in the elderly and remain non-negligible factors. Depression, in particular, is a common psychological disorder among older adults and continues to be one of the most prevalent and disabling biopsychosocial conditions in this population. A Chinese cross-sectional study provided evidence of the association between depressive symptoms and ADL disability [36]. There is also a strong association between depression and physical activity, with significant mental health benefits gained from being physically active, even at levels below public health recommendations [37]. This may explain why depression can affect disability progression through physical function measures such as handgrip strength and CS-5. Similarly, comorbidities and standing balance are associated with disability and are critical factors in the multifactorial process of disability [38,39]. Comorbidity, the coexistence of 2 or more chronic diseases in older adults is a well-documented risk factor for increased mortality, reduced quality of life, and functional decline, ultimately leading to disability [40]. As a consequence of managing multiple chronic conditions, polypharmacy, defined as the concurrent use of multiple medications, becomes increasingly common in older populations [41]. Polypharmacy has been associated with a higher risk of falls, frailty, cognitive impairment, and adverse drug interactions, further exacerbating health deterioration and disability [42]. However, in this study, medication use was categorized in the CHARLS questionnaire only as Chinese traditional medicine or Western modern medicine, without

detailed data on specific medications. This limitation prevented a comprehensive analysis of polypharmacy's impact. The decline in standing balance in older adults indicates decreased postural control and increased risk of falls, often seen in populations with sarcopenia and frailty, which eventually progresses to disability [43]. Overall, these findings suggest that we should consider the combined effects of mental health, management of multiple chronic diseases, and balance function when predicting disability and formulating prevention strategies.

In our study, we evaluated the treatment benefits of the model using DCA. Figure 6A shows the net benefit across different risk thresholds. At low thresholds (<0.2), the net benefit was high but gradually declined as the threshold increased, approaching zero around 0.6. The model outperformed both all-treatment and no-treatment strategies across most thresholds, demonstrating its clinical utility. In practice, selecting an appropriate threshold is critical for clinical decision-making and resource allocation. A low threshold is suitable for high-sensitivity scenarios, such as community screening for early intervention in older adults. A medium threshold balances sensitivity and specificity, making it ideal for resource-limited settings where the model can precisely identify high-risk individuals for targeted interventions. Many clinicians have used DCA to test various disease prediction models, such as those for 30-day mortality in MIMIC-III patients with sepsis-3, major adverse cardiovascular events in older patients, and hypertension risk in patients with type 2 diabetes mellitus [44-46].

- We evaluated the model's predictive efficacy at different risk thresholds using CIC. As shown in Figure 6C, at thresholds below 0.2, the model identifies over 500 high-risk patients, with approximately 200 actual events, resulting in a high false-positive rate and increased resource consumption. This range is suitable for early widespread screening when follow-up resources are available. At thresholds between 0.2 and 0.5, the number of high-risk patients identified aligns more closely with actual events, balancing sensitivity and specificity while improving cost-effectiveness. This range is ideal for resource-limited settings. At thresholds above 0.5, the number of high-risk patients decreases significantly, nearly matching actual events but potentially missing some high-risk cases. The CIC provides clinicians with a visual tool to balance sensitivity and specificity, optimizing disability prediction and intervention strategies in older adults. CIC is commonly used to evaluate the predictive accuracy and clinical value of clinical prediction model for various diseases [47-49]. However, they are rarely used to evaluate the clinical usefulness of disability prediction models in older adults.

The 9 variables selected through LASSO regression form a streamlined yet effective set of predictors that can be easily integrated into routine clinical practice. Specifically, the inclusion of sarcopenia and frailty-related features provides health care professionals with clear and actionable insights into the ability of older adults to live independently, enabling timely interventions to prevent disability. By focusing on these key variables, the model remains interpretable, reducing the risk of "black-box" complexity in clinical decision-making. To facilitate practical application, we developed a web-based application via the Streamlit platform that uses these 9 predictors to calculate the 5-year risk of disability for individual patients. This user-friendly tool allows clinicians to input patient-specific data and receive immediate risk assessments, integrating predictive analytics into clinical workflows and bridging complex data models with everyday decision-making. Future applications of this model can aid healthcare professionals in identifying individuals at high risk of disability and implementing early, targeted interventions. This approach has the potential to delay the onset of disability and improve the quality of life for older individuals.

Limitations

The limitations of this study include the selection of predictor variables. While the selected predictors are based on the best available evidence, other important variables, such as activity intensity, were not included due to high missing values. In addition, the model generalization and optimization is a limitation. The model performs well on the internal test set but lacks external validation due to the unavailability of a suitable external dataset. We plan to collect data from multicenter older care communities for external validation to further improve and optimize the model. As the CHARLS database updates, the model may need periodic updates to maintain accuracy and usefulness. Moreover, predicting long-term disability risk is challenging due to complex time interactions that may alter the risk trajectory.

Conclusions

Our research incorporates parameters aligned with the diagnostic criteria for sarcopenia and frailty. These physical function measures are combined with predictors from cognitive and psychological health dimensions, recognizing the complex interplay of physical capability, aging, and mental health in the development of disability. This approach enhances the model's precision and considers the need for the efficient identification of at-risk individuals and the optimization of medical resources in clinical practice. Consequently, the model provides a highly reliable disability prediction tool for older patients, health care workers, and policymakers. In the future, we will adjust the model based on updates to the CHARLS database to ensure its suitability for the older population in China. In addition, we will seek appropriate external databases for validation and promote the model's application across different ethnic groups.

Acknowledgments

We would like to thank the staff and participants of CHARLS. Publicly available datasets can be found at http://charls.pku.edu.cn/
.

This research was supported by a grant from the National Key R&D Program of China (No.: 2020YFC2005601).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model performance metrics at thresholds of 0.2 and 0.5 for the training and testing sets, respectively.
[DOC File, 32 KB - aging_v8i1e66723_app1.doc]

References

1. Foreman KJ, Marquez N, Dolgert A, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. Lancet 2018 Nov 10;392(10159):2052-2090. [doi: 10.1016/S0140-6736(18)31694-5] [Medline: 30340847]
2. Wang S. Spatial patterns and social-economic influential factors of population aging: A global assessment from 1990 to 2010. Soc Sci Med 2020 May;253:112963. [doi: 10.1016/j.socscimed.2020.112963] [Medline: 32289647]
3. Luo Y, Su B, Zheng X. Trends and challenges for population and health during population aging - China, 2015-2050. China CDC Wkly 2021 Jul 9;3(28):593-598. [doi: 10.46234/ccdcw2021.158] [Medline: 34594944]
4. Oppewal A, Hilgenkamp TIM, van Wijck R, Schoufour JD, Evenhuis HM. Physical fitness is predictive for a decline in the ability to perform instrumental activities of daily living in older adults with intellectual disabilities: Results of the HA-ID study. Res Dev Disabil 2015;41-42:76-85. [doi: 10.1016/j.ridd.2015.05.002] [Medline: 26079525]
5. Jung SW, Yoon JH, Lee W. Predictors for depressive symptoms by four types of disability. Sci Rep 2021 Sep 29;11(1):19371. [doi: 10.1038/s41598-021-98765-4] [Medline: 34588530]
6. Dodge HH, Kadowaki T, Hayakawa T, Yamakawa M, Sekikawa A, Ueshima H. Cognitive impairment as a strong predictor of incident disability in specific ADL-IADL tasks among community-dwelling elders: the azuchi study. Gerontologist 2005 Apr;45(2):222-230. [doi: 10.1093/geront/45.2.222] [Medline: 15799987]
7. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the china health and retirement longitudinal study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
8. Han Y, Wang S. Disability risk prediction model based on machine learning among Chinese healthy older adults: results from the China Health and Retirement Longitudinal Study. Front Public Health 2023;11(1271595):1271595. [doi: 10.3389/fpubh.2023.1271595] [Medline: 38026309]
9. Gill TM, Williams CS, Tinetti ME. Assessing risk for the onset of functional dependence among older adults: the role of physical performance. J American Geriatrics Society 1995 Jun;43(6):603-609 [FREE Full text] [doi: 10.1111/j.1532-5415.1995.tb07192.x]
10. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. J Aging Health 1993 May;5(2):179-193. [doi: 10.1177/089826439300500202] [Medline: 10125443]
11. Cheng ST, Chan ACM, Fung HH. Factorial structure of a short version of the Center for Epidemiologic Studies Depression Scale. Int J Geriat Psychiatry 2006 Apr;21(4):333-336. [doi: 10.1002/gps.1467]
12. Hu H, Cao Q, Shi Z, Lin W, Jiang H, Hou Y. Social support and depressive symptom disparity between urban and rural older adults in China. J Affect Disord 2018 Sep;237:104-111. [doi: 10.1016/j.jad.2018.04.076]
13. Mo S, Wang Y, Peng M, et al. Sex disparity in cognitive aging related to later-life exposure to ambient air pollution. Sci Total Environ 2023 Aug 15;886(163980):163980. [doi: 10.1016/j.scitotenv.2023.163980] [Medline: 37150467]
14. Hu Y, Peng W, Ren R, Wang Y, Wang G. Sarcopenia and mild cognitive impairment among elderly adults: The first longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2944-2952. [doi: 10.1002/jcsm.13081] [Medline: 36058563]
15. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556. [Medline: 22094840]
16. Yang M, Hu X, Wang H, Zhang L, Hao Q, Dong B. Sarcopenia predicts readmission and mortality in elderly patients in acute care wards: a prospective study. J Cachexia Sarcopenia Muscle 2017 Apr;8(2):251-258. [doi: 10.1002/jcsm.12163] [Medline: 27896949]
17. Chen J, Zhang F, Zhang Y, et al. Trajectories network analysis of chronic diseases among middle-aged and older adults: evidence from the China Health and Retirement Longitudinal Study (CHARLS). BMC Public Health 2024 Feb 22;24(1):559. [doi: 10.1186/s12889-024-17890-7] [Medline: 38389048]
18. Lu J. Predicted probability of disability. 2024. URL: https://disability-predictionmodel.streamlit.app [accessed 2025-01-15]
19. GMaCoD C. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1459-1544. [doi: 10.1016/S0140-6736(16)31012-1] [Medline: 27733281]

20. van der Gaag WH, Chiarotto A, Heymans MW, et al. Developing clinical prediction models for nonrecovery in older patients seeking care for back pain: the back complaints in the elders prospective cohort study. Pain 2021 Jun 1;162(6):1632-1640. [doi: 10.1097/j.pain.0000000000002161] [Medline: 33394879]
21. Nicholas M, Vlaeyen JWS, Rief W, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
22. Alene KA, Wangdi K, Colquhoun S, et al. Tuberculosis related disability: a systematic review and meta-analysis. BMC Med 2021 Sep 9;19(1):203. [doi: 10.1186/s12916-021-02063-9] [Medline: 34496845]
23. Watsford ML, Murphy AJ, Pine MJ. The effects of ageing on respiratory muscle function and performance in older adults. J Sci Med Sport 2007 Feb;10(1):36-44. [doi: 10.1016/j.jsams.2006.05.002]
24. Shin HI, Kim DK, Seo KM, Kang SH, Lee SY, Son S. Relation between respiratory muscle strength and skeletal muscle mass and hand grip strength in the healthy elderly. Ann Rehabil Med 2017 Aug;41(4):686-692. [doi: 10.5535/arm.2017.41.4.686] [Medline: 28971054]
25. Kera T, Kawai H, Hirano H, et al. Definition of respiratory sarcopenia with peak expiratory flow rate. J Am Med Dir Assoc 2019 Aug;20(8):1021-1025. [doi: 10.1016/j.jamda.2018.12.013] [Medline: 30737167]
26. Ohara DG, Pegorari MS, Oliveira Dos Santos NL, et al. Respiratory muscle strength as a discriminator of sarcopenia in community-dwelling elderly: a cross-sectional study. J Nutr Health Aging 2018;22(8):952-958. [doi: 10.1007/s12603-018-1079-4] [Medline: 30272099]
27. Vaz Fragoso CA, Beavers DP, Hankinson JL, et al. Respiratory impairment and dyspnea and their associations with physical inactivity and mobility in sedentary community-dwelling older persons. J Am Geriatr Soc 2014 Apr;62(4):622-628. [doi: 10.1111/jgs.12738] [Medline: 24635756]
28. Zhou Y, Li Y, Zhu X, Ma L. Medical and old-age care integration model and implementation of the integrated care of older people (ICOPE) in China: opportunities and challenges. J Nutr Health Aging 2021;25(6):720-723. [doi: 10.1007/s12603-021-1595-5] [Medline: 34179923]
29. Guido D, Leonardi M, Mellor-Marsá B, et al. Pain rates in general population for the period 1991-2015 and 10-years prediction: results from a multi-continent age-period-cohort analysis. J Headache Pain 2020 May 13;21(1):52. [doi: 10.1186/s10194-020-01108-3] [Medline: 32404046]
30. Dodds RM, Syddall HE, Cooper R, Kuh D, Cooper C, Sayer AA. Global variation in grip strength: a systematic review and meta-analysis of normative data. Age Ageing 2016 Mar;45(2):209-216. [doi: 10.1093/ageing/afv192] [Medline: 26790455]
31. Brooks JM, Titus AJ, Bruce ML, et al. Depression and handgrip strength among US. adults aged 60 years and older from NHANES 2011-2014. J Nutr Health Aging 2018;22(8):938-943. [doi: 10.1007/s12603-018-1041-5] [Medline: 30272097]
32. Coelho-Júnior HJ, Calvani R, Álvarez-Bustos A, et al. Physical performance and negative events in very old adults: a longitudinal study examining the ilSIRENTE cohort. Aging Clin Exp Res 2024;36(1):38345698. [doi: 10.1007/s40520-024-02693-y]
33. Garrett SL, Kennedy RE, Sawyer P, Williams CP, Brown CJ, Allman RM. Association between executive dysfunction and instrumental activities of daily living: racial and ethnic differences among community-dwelling older Adults in the Southeastern US. J Natl Med Assoc 2019 Jun;111(3):320-327. [doi: 10.1016/j.jnma.2018.11.002] [Medline: 30527966]
34. Sun Q, Jiang N, Lu N, Lou VWQ. Bidirectional relationship between cognitive function and loss hierarchy of activities of daily living among older adults with disabilities in urban China: a cross-lagged analysis. BMJ Open 2022 Sep;12(9):e057211. [doi: 10.1136/bmjopen-2021-057211]
35. Ali N, Tian H, Thabane L, et al. The effects of dual-task training on cognitive and physical functions in older adults with cognitive impairment; a systematic review and meta-analysis. J Prev Alzheimers Dis 2022;9(2):359-370. [doi: 10.14283/jpad.2022.16] [Medline: 35543010]
36. Feng Z, Li Q, Zhou L, Chen Z, Yin W. The relationship between depressive symptoms and activity of daily living disability among the elderly: results from the China Health and Retirement Longitudinal Study (CHARLS). Public Health (Fairfax) 2021 Sep;198:75-81. [doi: 10.1016/j.puhe.2021.06.023]
37. Pearce M, Garcia L, Abbas A, et al. Association between physical activity and risk of depression: a systematic review and meta-analysis. JAMA Psychiatry 2022 Jun 1;79(6):550-559. [doi: 10.1001/jamapsychiatry.2022.0609] [Medline: 35416941]
38. Vargese SS, Jylhä M, Raitanen J, Enroth L, Halonen P, Aaltonen M. Dementia-related disability in the population aged 90 years and over: differences over time and the role of comorbidity in the vitality 90 + study. BMC Geriatr 2023;23(1):37149593. [doi: 10.1186/s12877-023-03980-5]
39. Boyle PA, Wilson RS, Buchman AS, et al. Lower extremity motor function and disability in mild cognitive impairment. Exp Aging Res 2007;33(3):355-371. [doi: 10.1080/03610730701319210] [Medline: 17497375]
40. Yarnall AJ, Sayer AA, Clegg A, Rockwood K, Parker S, Hindle JV. New horizons in multimorbidity in older adults. Age Ageing 2017 Nov 1;46(6):882-888. [doi: 10.1093/ageing/afx150] [Medline: 28064173]
41. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. BMC Geriatr 2017 Oct 10;17(1):230. [doi: 10.1186/s12877-017-0621-2] [Medline: 29017448]

42. van Dam CS, Labuschagne HA, van Keulen K, et al. Polypharmacy, comorbidity and frailty: a complex interplay in older patients at the emergency department. Eur Geriatr Med 2022 Aug;13(4):849-857. [doi: 10.1007/s41999-022-00664-y] [Medline: 35723840]
43. Billot M, Calvani R, Urtamo A, et al. Preserving mobility in older adults with physical frailty and sarcopenia: opportunities, challenges, and recommendations for physical activity interventions. Clin Interv Aging 2020;15:1675-1690. [doi: 10.2147/CIA.S253535] [Medline: 32982201]
44. Hou N, Li M, He L, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. J Transl Med 2020 Dec 7;18(1):462. [doi: 10.1186/s12967-020-02620-5] [Medline: 33287854]
45. Zhang K, Liu C, Sha X, et al. Development and validation of a prediction model to predict major adverse cardiovascular events in elderly patients undergoing noncardiac surgery: a retrospective cohort study. Atherosclerosis 2023 Jul;376:71-79. [doi: 10.1016/j.atherosclerosis.2023.06.008] [Medline: 37315395]
46. Yang J, Wang X, Jiang S. Development and validation of a nomogram model for individualized prediction of hypertension risk in patients with type 2 diabetes mellitus. Sci Rep 2023 Jan 23;13(1):1298. [doi: 10.1038/s41598-023-28059-4] [Medline: 36690699]
47. Shen J, Xiao Z, Wang X, Zhao Y. A nomogram clinical prediction model for predicting urinary infection stones: development and validation in A retrospective study. World J Urol 2024;42(1):38573354. [doi: 10.1007/s00345-024-04904-7]
48. He M, Li C, Kang Y, Zuo Y, Duo L, Tang W. Clinical predictive model for the 1-year remission probability of IgA vasculitis nephritis. Int Immunopharmacol 2021 Dec;101:108341. [doi: 10.1016/j.intimp.2021.108341]
49. Hu P, Xu Y, Liu Y, et al. An externally validated dynamic nomogram for predicting unfavorable prognosis in patients with aneurysmal subarachnoid hemorrhage. Front Neurol 2021;12:683051. [doi: 10.3389/fneur.2021.683051] [Medline: 34512505]

Abbreviations

ADL: activities of daily living
ASM: appendicular skeletal muscle mass
AUC: area under the curve
CES-D: Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
CIC: clinical impact curve
CS-5: chair stand test
DCA: decision curve analysis
IADL: instrumental activities of daily living
LDL: low-density lipoprotein
ROC: receiver operating characteristic
SHAP: SHapley Additive exPlanations
TICS: telephone interview for cognitive status

Edited by P Jansons; submitted 21.09.24; peer-reviewed by J Hu, J Yue, L Ma, Y Qiu; revised version received 26.02.25; accepted 13.03.25; published 17.04.25.

Please cite as:
Chu J, Li Y, Wang X, Xu Q, Xu Z
Development of a Longitudinal Model for Disability Prediction in Older Adults in China: Analysis of CHARLS Data (2015-2020)
JMIR Aging 2025;8:e66723
URL: https://aging.jmir.org/2025/1/e66723
doi:10.2196/66723

Real-World Insights Into Dementia Diagnosis Trajectory and Clinical Practice Patterns Unveiled by Natural Language Processing: Development and Usability Study

Hunki Paek[1*], PhD; Richard H Fortinsky[2*], PhD; Kyeryoung Lee[1], PhD; Liang-Chin Huang[1], PhD; Yazeed S Maghaydah[2], MD; George A Kuchel[2], MD; Xiaoyan Wang[3,4], PhD

[1]IMO Health, Rosemont, IL, United States

[2]UConn Center on Aging, University of Connecticut School of Medicine, Farmington, CT, United States

[3]Center for Quantitative Medicine, University of Connecticut School of Medicine, 195 Farmington Ave, Farmington, CT, United States

[4]Department of Health Policy and Management, Tulane University, New Orleans, LA, United States

[*]these authors contributed equally

Corresponding Author:
Xiaoyan Wang, PhD
Center for Quantitative Medicine, University of Connecticut School of Medicine, 195 Farmington Ave, Farmington, CT, United States

Abstract

Background: Understanding the dementia disease trajectory and clinical practice patterns in outpatient settings is vital for effective management. Knowledge about the path from initial memory loss complaints to dementia diagnosis remains limited.

Objective: This study aims to (1) determine the time intervals between initial memory loss complaints and dementia diagnosis in outpatient care, (2) assess the proportion of patients receiving cognition-enhancing medication prior to dementia diagnosis, and (3) identify patient and provider characteristics that influence the time between memory complaints and diagnosis and the prescription of cognition-enhancing medication.

Methods: This retrospective cohort study used a large outpatient electronic health record (EHR) database from the University of Connecticut Health Center, covering 2010 - 2018, with a cohort of 581 outpatients. We used a customized deep learning–based natural language processing (NLP) pipeline to extract clinical information from EHR data, focusing on cognition-related symptoms, primary caregiver relation, and medication usage. We applied descriptive statistics, linear, and logistic regression for analysis.

Results: The NLP pipeline showed precision, recall, and F_1-scores of 0.97, 0.93, and 0.95, respectively. The median time from the first memory loss complaint to dementia diagnosis was 342 (IQR 200-675) days. Factors such as the location of initial complaints and diagnosis and primary caregiver relationships significantly affected this interval. Around 25.1% (146/581) of patients were prescribed cognition-enhancing medication before diagnosis, with the number of complaints influencing medication usage.

Conclusions: Our NLP-guided analysis provided insights into the clinical pathways from memory complaints to dementia diagnosis and medication practices, which can enhance patient care and decision-making in outpatient settings.

(JMIR Aging 2025;8:e65221) doi:10.2196/65221

KEYWORDS

dementia; memory loss; memory; cognitive; Alzheimer disease; natural language processing; NLP; deep learning; machine learning; real-world insights; electronic health records; EHR; cohort; diagnosis; diagnostic; trajectory; pattern; prognosis; geriatric; older adults; aging

Introduction

The rising prevalence of dementia, driven by an aging population, presents a profound concern for society [1-4], placing substantial burdens on individuals and imposing high financial costs [5-7]. Despite these challenges, no curative treatments are currently available [8,9], highlighting the critical need for early detection of prodromal symptoms of dementia, such as mild cognitive decline [10-13], and timely diagnosis. Early intervention can delay disease progression or alter the trajectory toward dementia [9,14,15]. However, dementia and its associated symptoms are frequently underreported and underdiagnosed in clinical practice [16,17]. As the condition progresses, patients commonly experience increased memory loss, deteriorating cognitive ability, heightened confusion, and changes in personality like agitation. The conversion from mild

cognitive impairment to Alzheimers disease has been explored using patient health data [18].

Electronic health records (EHRs) offer a valuable resource for enhancing the detection and management of disease by providing comprehensive data on patient health and history [19-23]. However, much of the nuanced patient information is embedded within unstructured clinical notes and is not accessible through structured data. Natural language processing (NLP), a subfield of artificial intelligence that enables computers to understand, interpret, and generate human language, holds promise for extracting meaningful information from vast and complex free-text EHRs [24-27]. NLP has been instrumental in automatically extracting clinical information in various medical domains, including geriatric care [28-35]. For instance, Kharrazi et al [36] showcased higher rates of geriatric syndrome extraction from unstructured EHR using NLP compared to relying solely on claim data or structured EHR data. Studies have successfully extracted cognitive status and measurement scores [37,38], as well as lifestyle exposures and discourse production for Alzheimers disease [39], from clinical documentation using NLP. Additionally, multiple studies have applied NLP methods to extract neuropsychiatric symptoms and cognitive or function impairment information [40-43]. State-of-the-art models, such as pretrained Bidirectional Encoder Representation from Transformer (BERT), have been applied in clinical settings for tasks like detecting inpatient falls [44] and classifying dementia risk [45] using clinical notes. While previous research has made significant strides in the earlier detection of cognitive decline using EHR, most studies have focused on extracting symptoms or cognitive measurement scores rather than other clinical features that affect disease progression, such as the relationship of the primary caregivers with patients.

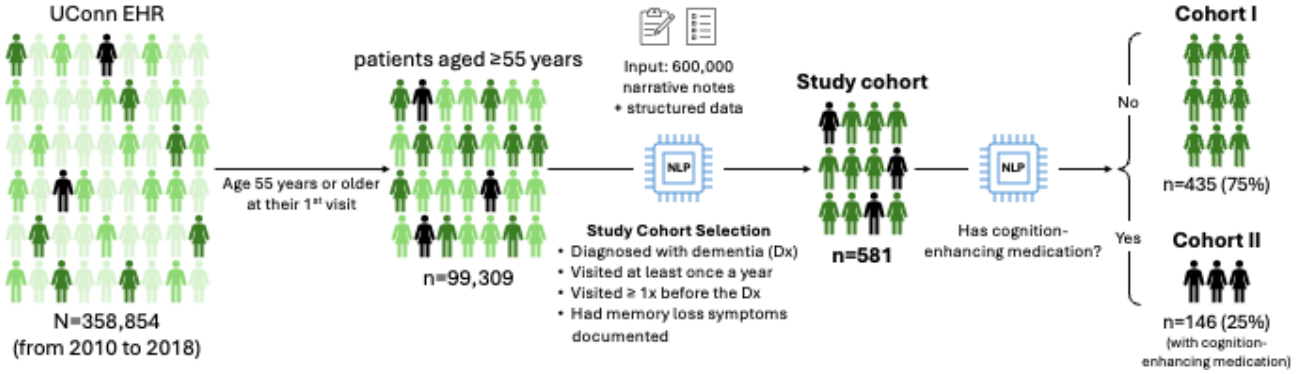
We aimed to investigate the time interval from initial memory loss complaints to dementia diagnosis and explore the association between various clinical features, including the family primary caregiver relationship, using real-world outpatient clinical notes. Additionally, we aimed to analyze the pattern of cognition-enhancing medication prescriptions before diagnosis. To achieve this, we developed a customized NLP pipeline using deep learning techniques, based on a prodromal dementia symptom ontology that we established.

Methods

Study Cohorts

This retrospective study used data from the UConn Health Center between 2010 and 2018. The use of longitudinal EHR data allowed us to track all patients' clinical information, including demographic characteristics, diagnoses, measurements, medications, and signs. The study cohort was defined as patients who met the following criteria: (1) received a dementia diagnosis, (2) had at least one outpatient visit per year, (3) had at least one visit before the dementia diagnosis, and (4) had documented memory loss–related symptoms (eg, memory loss, confusion, cognition impairment, trouble remembering, not recalling, forgetting, and blackout) in the EHR. Dementia was defined based on the presence of 3 or more *ICD* (*International Classification of Diseases*) codes used for dementia phenotyping in our study (as detailed in Multimedia Appendix 1) and dementia documentation in their clinical notes. We analyzed demographic and clinical characteristics from structured EHR data, including insurance details, the initial location (medical unit) of memory loss complaints, and the location of the first dementia diagnosis. These locations encompass various settings within this health care system, such as geriatric medicine, internal medicine, and neurology outpatient clinics. NLP was used to extract symptoms and primary caregiver (family supporter) relationship information from clinical notes. Both diagnosis and cognition-improving medication information were extracted from both structured and unstructured data. Figure 1 provides an overview of the selection process of the study cohort and the information extraction process.

Figure 1. The workflow for the study cohort selection and information extraction. EHR: electronic health record; NLP: natural language processing.



Building a Deep Learning–Based NLP Pipeline

A framework was developed to curate the prodromal dementia symptoms comprising four stages: (1) preprocessing and query expansion, (2) ontology construction and annotation, (3) NLP model development, and (4) system evaluation.

Preprocessing and Query Expansion

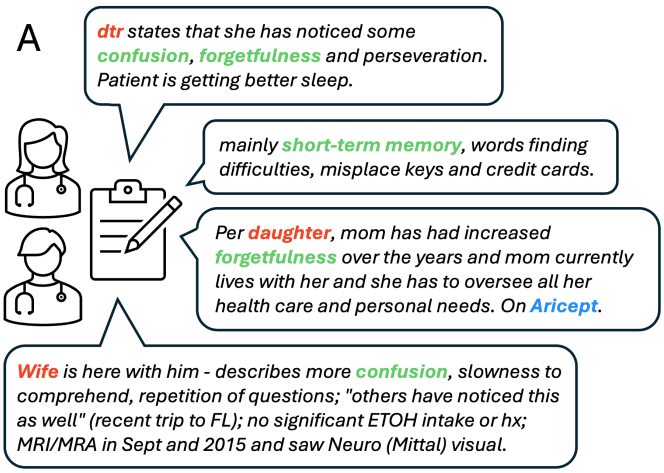
In this stage, a query was expanded to extract and identify a broad set of patient notes with documented memory loss symptoms. A list of seed terms was obtained through a manual survey of the literature and a clinical note review by a clinical researcher and domain expert. A bigram word2vec algorithm [46] was used to identify additional significant terms potentially related to memory loss symptoms to ensure the encapsulation

of an expansive cohort. The expanded, rule-based query terms provided in Multimedia Appendix 2 were subsequently applied to extract the relevant patient notes for NLP modeling.

Ontology Construction and Annotation

This stage involves the simulation of an expert's knowledge and understanding of the free text. A prodromal dementia symptom ontology was built based on the physician's opinion, a comprehensive literature review, and a clinical note review. The ontology includes 9 entities and 9 relations. The 9 entities are "memory loss symptom (Sx)," "dementia diagnosis (Dx)," "temporal," "duration," "status change: worse," "other symptoms," "cognitive test result," "caregiver relation," and "cognition enhancing medication (Rx)." The 9 relations are "has complaint date," "has diagnosis date," "has other symptom information," "has status change information," "has duration information," "has test information," "has caregiver information," "has treatment information," and "has effects" as depicted in Figure 2A. Two independent annotators manually annotated notes using Clinical Language Annotation, Modeling, and Processing (CLAMP) [47], an NLP toolkit, guided by the constructed ontology. Figure 2B shows an example note with entities and relations annotated.

Figure 2. The ontology of memory loss and the annotated sample note. (A) The ontology of prodromal dementia symptoms. In total, 9 entities and 9 semantic relations between entities were defined in the ontology. (B) A sample note that has been deidentified and annotated with prodromal dementia symptoms. In this note, the following entities have been annotated and related to each other: symptoms, relation, medication, duration, status change, test, and temporal. MRI/MRA: magnetic resonance imaging/magnetic resonance angiography, dtr:daughter, ETOH: ethanol, FL: florida, hx: history.



B

Entity	Standardized concept values	Original strings
Symptom	Memory loss	confusion
Symptom	Memory loss	forgetfulness
Symptom	Memory loss	short-term memory
Medication	Donepezil	Aricept
Relation	Daughter	drt
Relation	Daughter	daughter
Relation	Wife	Wife

NLP Model Development

The annotated notes (n=815) obtained in the previous stage were used for NLP model training. These notes were randomly split into a training set (80%, n=652, of annotated notes) and an independent validation set (the remaining 20%, n=163). The manual annotation and training processes were iteratively performed with additional manually annotated notes to enhance model performance until the model achieved an F_1-score of >0.8. For model training, a multilayer deep learning architecture was adopted, which involved transforming the text into sequential vectors of characterization through the embedding step. The vectors were then fed into a Bidirectional Long Short-Term Memory (BiLSTM), a text classification architecture based on artificial neural networks, for pattern recognition in both forward and backward directions. Finally, the patterns were sent to the next layer of a Conditional Random Field (CRF) model for prediction probability computation. BiLSTM-CRF architectures are widely used in clinical NLP tasks, demonstrating robustness in recognizing entities in sequential data like clinical notes [48] and effectively leveraging moderate-sized datasets with lower computational resource requirements.

NLP Pipeline Evaluation

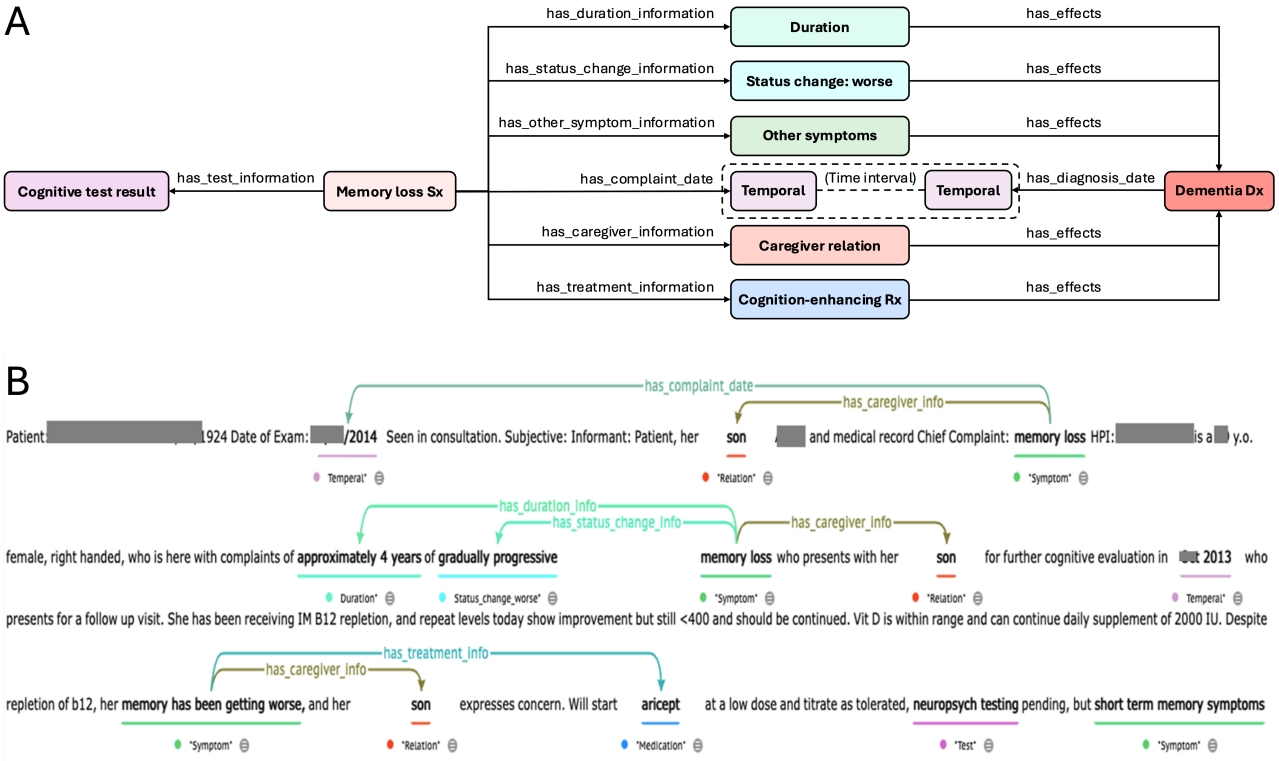
The performance of the pipeline was evaluated in the validation set through precision (positive predictive value [PPV]), recall (sensitivity), and F_1-score (a balanced score between false positives [FPs] and false negatives [FNs]). Recall was calculated as the ratio of the number of entities that were identified by the pipeline over the total number of the corresponding entities in the manually annotated gold standard (ie, true positive [TP]/[TP+FN]). Precision was measured as the ratio of the number of distinct entities returned by the correct pipeline according to the gold standard divided by the total number of entities found by our pipeline (ie, TP/[TP+FP]). F_1-score was calculated as the harmonic mean of PPV and sensitivity (ie, 2 × PPV × sensitivity / [PPV + sensitivity]).

Standardization of Concept Values Leveraging NLP

Clinical notes contain abundant information but are often heterogeneous in form. To enable use case analysis, these heterogeneous entities needed to be standardized. Figure 3A illustrates the various forms in which cognition-related concepts like forgetfulness, memory loss, short-term memory, and confusion were documented in the clinical notes. Abbreviations (eg, dtr for daughter) and mixed use of brand names and generic names of the same drugs (eg, Aricept and donepezil) were commonly found. Figure 3B provides examples of standardized concept values obtained through the NLP process. The cognition-enhancing medications discussed include donepezil, memantine, rivastigmine, and galantamine. For the analysis, the son, son-in-law, daughter-in-law, grandson, and granddaughter were classified as other adult children, while the

nephew, niece, cousin, brother, and sister were classified as other family support.

Figure 3. Clinical note standardization via NLP. (A) The diverse expressions used in clinical notes depict concepts such as memory loss symptoms, primary caregiver relationships, and cognition-enhancing medications. (B) A sample of the standardized output generated by the NLP process, representing the standardized values of the various original expressions. Dx: diagnosis; HPI: history of present illness; NLP: natural language processing; Rx: medication; Sx: symptom.



Characterization of Dementia Cohorts and Longitudinal Analysis of Dementia Trajectory

The collection of clinical information of patients was accomplished through the extraction of both structured and unstructured EHR data. The statistical analysis was performed using R software (R Core Team). A logistic regression model was used to assess the relationship between medication prescription history and other factors, such as memory loss complaints and the primary caregiver relation information, by calculating the odds ratio (OR) and 95% CI. A total of 149 patients who had memory loss complaints and diagnoses recorded on the same day were excluded from this study.

Ethical Considerations

This study involved no interaction with patients as it is a retrospective cohort study that used a deidentified dataset. As such, it was deemed as exempt from institutional review board approval.

Results

Building a Memory Loss NLP Pipeline

The performance of our memory loss NLP pipeline was evaluated using precision, recall (sensitivity), and F_1-score metrics in the validation set, with detailed results presented in Multimedia Appendix 3. Our system achieved high scores across all semantic types, including “memory loss symptom,” “dementia diagnosis,” “duration,” “primary caregiver,” and “status change” (overall precision, recall [sensitivity], and F_1-scores of 0.97, 0.93, and 0.95, respectively). For example, the precision for “memory loss symptoms” was 0.97, indicating that 97% of “memory loss symptoms” were identified by our NLP system in patients’ clinical notes, as verified against a manually curated gold standard. The recall (sensitivity) of 0.93 implies that our system correctly identified 93% of actual memory loss cases, with only 7% missed. Multimedia Appendix 3 provides comprehensive details for various semantic types and relations, enabling a thorough understanding of the NLP pipeline’s performance.

Clinical Characterization of Study Cohorts

Cohort Identification

An average of 358,854 patients visited the UConn Health system for outpatient care between 2010 and 2018, with the number of visits increasing annually from 224,488 in 2010 to 1,024,349 in 2017. There were 8686 patient visits in 2018 until the time point we collected data. Out of these patients, 99,039 patients were aged 55 years or older at their first visit. From more than 600,000 narrative records and coded data of those patients, we identified 730 patients who had at least one outpatient visit per year, at least one visit before their dementia diagnosis, and documented memory loss symptoms in their clinical notes. Of these, 149 (20.4%) patients reported memory loss complaints and were diagnosed with dementia on the same day, while 581 (79.6%) patients had at least a 1-day gap between the complaint and diagnosis. For the following analysis, 581 patients with

memory loss symptoms documented at different days were included in study cohorts (Figure 1).

Cohort Demographics

The demographic characteristics, primary insurance, and medication information of study cohorts are summarized in Table 1. Most cohort members were non-Hispanic White individuals (509/581, 87.6%) and female (381/581, 65.6%), with an age distribution of over 85 years (315/581, 54.2%), 75 - 84 years (171/581, 29.4%), 65-74 years (63/581, 10.8%), and under 65 years (32/581, 5.5%).

Table . Demographic characteristics, primary insurance, and medication information of study cohorts (N=581).

Characteristics		Values, n (%)
Age (years)		
	<65	32 (5.5)
	65 - 74	63 (10.8)
	75 - 84	171 (29.4)
	85+	315 (54.2)
Race		
	White	509 (87.6)
	African American	33 (5.7)
	Others	39 (6.7)
Sex		
	Female	381 (65.6)
	Male	200 (34.4)
Primary caregiver (family supporter) relation (n=291)		
	Husband	26 (8.9)
	Wife	39 (13.4)
	Daughter	135 (46.4)
	Son	50 (17.2)
	Other family members	41 (14.1)
	Missing	290 (49.9)
Prior medication before the first diagnosis of dementia		146 (25.1)

Primary Caregiver and Medication Information

Out of 291 patients with primary caregiver relation information, adult children were the main caregivers (n=185, 63.6%, with n=135, 46.4%, being daughters and n=50, 17.2%, being sons), followed by spouses (n=65, 22.3%, with n=39, 13.4%, being wives and n=26, 8.9%, being husbands). Other family members (eg, nephew) made up 14.1% (41/291) of the cohort. Out of 581 patients, 146 (25.1%) patients had been prescribed cognition-improving medications prior to the dementia diagnosis.

Outpatient Care Locations

Next, we investigated the outpatient care locations where the first memory loss complaints were reported and where dementia was diagnosed. Geriatrics is the most frequent location for both the first memory loss complaints made (308/581, 53%) and the diagnosis of dementia (350/581, 60.2%), followed by primary care (185/581, 31.8%, and 163/581, 28.1%, respectively) and neurology (39/581, 6.7%, and 61/581, 10.5%, respectively). The majority of the cohort was covered by Medicare (354/581, 60.9%) or Medicaid (72/581, 12.4%) as primary insurance, while 26.2% (152/581) had commercial insurance. Only a small percentage of patients (2/581, 0.3%) had no insurance coverage (Multimedia Appendix 4).

Distribution of Time Intervals Between Cognitive Symptom Complaints and Dementia Diagnosis and the Number of Complaints

Time Interval

The median time interval between the first memory loss complaints and dementia diagnosis was 342 days, ranging from a minimum of 1 day to a maximum of 1458 days in our study cohort (n=581) (Multimedia Appendix 5).

Health Care Use

Additionally, the number of complaints made before being diagnosed was analyzed, with a median of 3 complaints, ranging from a minimum of 1 complaint to a maximum of 18 complaints (Multimedia Appendix 5).

Association Analysis for the Earlier Dementia Diagnosis

We aimed to identify the clinical features that are associated with earlier diagnosis of dementia from the first memory loss

complaints. Results indicated that the location of the first complaints made and the diagnosis, as well as the relation of the primary caregiver, were significantly associated with earlier diagnosis of dementia. Patients who made complaints in geriatrics (−141 days, $P<.001$, χ^2 test) or neurology (−158 days, $P=.02$) were diagnosed with dementia earlier compared to those who made complaints in primary care. Furthermore, patients diagnosed with dementia in geriatrics had a shorter interval of 152.9 days ($P<.001$) compared to those diagnosed in primary care. Additionally, having a wife or a daughter as a primary caregiver was associated with an earlier diagnosis of dementia, with a shorter interval of 249.6 days ($P=.01$) and 176.8 days ($P=.04$), respectively, compared to those who had a husband as a primary caregiver. However, factors such as age or insurance types were not found to have a significant impact on earlier diagnosis (Table 2).

Table . Statistical analysis of time intervals between the first complaints and dementia diagnosis.

Features		Estimate (95% CI)	*P* value (χ^2 test)
Age (years)			
	<65	109.8 (–26.7 to 246.3)	.12
	65 - 74	–61.6 (–163.2 to 40)	.24
	75 - 84	0.08 (–69.4 to 69.6)	≥.99
	85+	Ref[a]	—[b]
Primary insurance			
	No insurance	36.8 (–25 to 98.6)	.24
	Medicaid	5.8 (–255.5 to 267.1)	.97
	Medicare	36.8 (–11.9 to 85.5)	.14
	Commercial	Ref	—
The location of the first memory loss complaint			
	Geriatrics	–141 (–212 to –70)	<.001
	Neurology	–158 (–286.2 to –29.8)	.02
	Primary care	Ref	—
	Other	81 (–35.1 to 197.1)	.17
The location of the first diagnosis of dementia			
	Geriatrics	–152.9 (–244 to –61.8)	<.001
	Neurology	–82.2 (–191.4 to 27)	.14
	Primary care	Ref	—
	Other	42 (–225.4 to 309.4)	.76
Primary caregiver (family supporter) relation			
	Wife	–249.6 (–437.2 to –62)	.009
	Daughter	–176.8 (–346.1 to –7.5)	.04
	Other adult children	–127.4 (–256.9 to 2.1)	.05
	Other family support	–257.5 (–577.1 to 62.1)	.11
	Husband	Ref	—

[a]Reference.

[b]Not applicable.

Association Analysis for the Medication Usage

Medication was prescribed in 25.1% (146/581) of patients before dementia diagnosis. We next analyzed factors associated with the usage of cognition-enhancing medication before the diagnosis of dementia after the 1st complaints of memory loss. The only factor that was significantly associated with medication usage was the total number of memory loss complaints made; each additional memory complaint was associated with a 15% greater likelihood that cognition-enhancing medications were prescribed (OR 1.148, 95% CI 1.027 - 1.283; Table 3).

Table . An analysis of the factors associated with the usage of medication before the diagnosis of dementia.

Features		OR[a] (95% CI)
Age (years)		
	<65	3.827 (0.403 - 23.32)
	65 - 74	1.251 (0.507 - 3.727)
	75 - 84	1.127 (0.615 - 2.445)
	85+	Ref[b]
The location of the first memory loss complaint		
	Geriatrics	1.477 (0.551 - 3.959)
	Neurology	2.124 (0.449 - 10.05)
	Primary care	Ref
	Other	0.331 (0.103 - 1.058)
The location of the first diagnosis		
	Geriatrics	0.489 (0.172 - 1.39)
	Neurology	0.65 (0.182 - 2.319)
	Primary care	Ref
Family support		
	Wife	4.367 (0.85 - 22.447)
	Daughter	1.831 (0.263 - 12.74)
	Other adult children	1.609 (0.538 - 4.816)
	Other family support	1.033 (0.276 - 3.871)
	Husband	Ref
Total number of memory loss complaints before the diagnosis of dementia		1.148 (1.027 - 1.283)

[a]OR: odds ratio.

[b]Reference.

Discussion

Principal Findings

We developed a high-performance deep learning–based NLP algorithm on an EHR dataset of dementia patients to delve into the real-world trajectory of dementia, starting from initial memory loss complaints to dementia diagnosis. Our investigation focused on the time interval from the first memory loss complaints to dementia diagnosis, the proportion of prescribed cognition-enhancing medication before diagnosis during this trajectory, and the clinical characteristics associated with these features.

We found that 20.4% (149/730) of patients had same-day documentation of memory loss complaints and dementia diagnosis. Among the remaining 79.5% (580/730) of patients with at least a 1-day gap between complaints and diagnosis, over half of the patients received a dementia diagnosis within a year of their initial memory loss complaints, with a median time of 342 days. The location of the first complaint and diagnosis and the relationship with the primary caregiver emerged as influential factors in achieving an earlier diagnosis. Notably, patients who initiated complaints or were diagnosed in geriatrics or neurology received earlier diagnoses compared to those in primary care. This underscores the important role of the initial complaint's location and the dementia diagnosis's setting in the early detection and management of dementia. Our findings align with previous research indicating missed and delayed diagnoses in primary care [49]. Geriatricians and neurologists possess significantly more expertise and practical experience in diagnosing dementia and prescribing these meds than most primary care doctors. To enhance early detection in primary care, it is crucial to train primary care providers to recognize the nuances of dementia symptoms and to appreciate the importance of thorough assessments during initial consultations. This approach could significantly reduce delays in diagnosis. Furthermore, improving caregiver education regarding the signs and symptoms of dementia is essential, as it can lead to earlier recognition of concerns and encourage timely visits to health care professionals. Additionally, understanding the factors that lead patients to receive care in a geriatric or neurological department rather than primary care would be an important question for further investigation. Similar to the previous study that identified dementia severity and marital status as independent predictors of receiving a clinical cognitive evaluation [50], other factors such as more complex medical needs (eg, multiple chronic conditions and polypharmacy), severe function decline, and the primary caregiver's educational level or relationship with the patient

could be associated with visits to geriatric or neurologic departments.

Remarkably, we found that patients with a wife or daughter as their primary caregiver were diagnosed earlier and more frequently used cognitive-improving medication before the dementia diagnosis. This emphasizes the vital role of primary caregivers in the diagnosis and treatment of dementia patients. Mahmoudi et al [35] previously emphasized the importance of extracting caregiver information in dementia patient notes and developed the rule-based NLP algorithm to identify caregiver availability. In our work, we extended this by also extracting family-caregiver relationships with patients and analyzing their impact on the early diagnosis of dementia. Subsequent research should explore the underlying mechanisms and factors of caregivers in this context such as the association between the relationship of primary caregivers and visits to geriatric or neurologic departments. Contrary to expectations, our study revealed that age and insurance were not associated with earlier diagnoses.

Surprisingly, the total number of memory loss complaints emerged as the sole factor significantly linked to medication usage, with other factors showing no significant association. The correlation between increased medication prescriptions and additional memory loss complaints highlights health care providers' responsiveness to escalating symptoms and underscores the importance of proactive monitoring of cognitive symptoms. This finding aligns with the Alzheimer's Association guidelines for early identification and treatment of Alzheimers disease, particularly in its initial stages. By recognizing and documenting memory loss complaints, clinicians can better initiate appropriate therapeutic interventions, particularly those aimed at altering disease progression. Additionally, our findings emphasize the need for enhanced caregiver education on the importance of reporting memory loss and other cognitive changes.

We demonstrated that extracting cognitive symptom-related terms from longitudinally documented patient notes before dementia diagnosis could be an alternative approach to analyzing documented cognitive measurement scores during patient visits, potentially aiding in identifying dementia patients. Previous studies have highlighted a significant lack of such documentation in clinical notes [37,41,51]. For instance, Harding et al [51] found that cognitive measurement scores were rarely available in their cohorts when establishing the algorithm for identifying dementia patients in EHR. Similarly, Maserejian et al [37] demonstrated a low percentage of dementia (11%) or Alzheimers disease (24%) in patients with cognitive measurement scores such as Mini Mental State Examination (MMSE), a recall test, a clock drawing, Montreal Cognitive Assessment (MoCA), Mini-Cog, or Saint Louis University Mental Status (SLUMS) documented and suggested prompts of cognitive measurement. McCoy et al [41] attempted to extract cognitive symptom-related terms (eg, impulsive, forgetful, cognitive, and memory) and converted them into scores, given the issue of reliability and scalability of the cognitive measurement test. Consistently, the proportion of patient notes with cognitive test names, including MMSE, SLUMS, MoCA, Mini-Cog, clock drawing, trail making, Boston naming test, and Wisconsin card sorting test, was very low in our study, so these were not used in further analysis.

Our NLP approach in automatically identifying cognitive symptom-related terms and primary caregivers, and systematically analyzing these factors along with other structured data, enhanced our understanding of dementia progression and management. This approach provides a practical and scalable method for identifying cognitive impairment, especially when traditional cognitive measurement scores are lacking. Further exploration using this NLP method could significantly advance the field, providing deeper insights and more effective interventions for dementia care. Moreover, our study findings align with the Alzheimer's Association's recommendations for using simple practical tests like Mini-Cog or General Practitioner Assessment of Cognition (GPCOG), developed by a group of clinical dementia experts, during annual visits, particularly when symptoms are reported by patients or caregivers, thereby supporting current clinical practice in dementia care.

Strengths and Limitations

Our study has several strengths. First, our work presents a novel approach to understanding clinical practices in dementia by examining the time interval from the symptom complaints to the diagnosis of dementia using real-world data. Second, our study highlights how the relationship between primary caregivers and patients and the location (medical units: geriatrics, neurology, and primary care) of complaints made may influence the time to diagnosis. While existing ontologies provide comprehensive medical concepts, they often lack the specificity necessary to accurately capture the relationships between concepts. By defining the relationships between entities such as "has date" and "has caregiver information" in the ontology, our study could provide deeper insights into the clinical features that affect the time interval from the symptom presentation to dementia diagnosis in real-world contexts. Third, our study developed and validated a customized NLP model to be used to predict an outcome in a clinical setting using EHRs.

Several limitations of this study should be considered. First, the study relied on EHR data from a single health care system, which may limit the generalizability of the findings to other populations and health care settings. This system-specific reliance may introduce potential biases related to local clinical practices, documentation standards, and patient demographics, which could affect the study's findings if applied to broader or more diverse health care environments. Additionally, the patient population in our dataset was predominantly White, female, and older patients above 85 years with Medicare or Medicaid insurance, which may further limit the generalizability of the findings to other demographic groups, including younger patients, males, and individuals from diverse racial backgrounds. Future studies should aim to include a more diverse patient population in terms of age, race, and insurance type across multiple health care systems to validate and potentially broaden the applicability of the findings.

Another limitation is the potential for loss of follow-up within the EHR data, as patients with less frequent or inconsistent visits may have different clinical trajectories. We preselected patients

who had been diagnosed with dementia, had at least one documented memory loss complaint before diagnosis, and visited the health care system at least once per year to reduce the likelihood of significant follow-up loss. Nonetheless, variability in follow-up could still influence our results. Furthermore, potential biases within EHR documentation may impact the findings. For example, memory complaints may be underreported or inconsistently documented, depending on clinician practices and the completeness of note-taking. This variability could affect the accuracy of data extraction and the insights derived from the patient journey. Lastly, the study did not account for the potential impact of other medical conditions on the result. Incorporating objective measures of cognitive decline and other co-occurring neuropsychiatric symptoms could enhance the assessment of dementia.

Conclusions

Our study highlights the importance of the location of initial memory loss complaints, the location of the dementia diagnosis, and the role of the primary caregiver in the early diagnosis and treatment of dementia patients. By analyzing complex clinical dementia care practice patterns within a real-world setting on a large scale using NLP, our exploratory analysis demonstrates the potential of advanced analytical techniques in achieving earlier and more accurate diagnoses of dementia.

Acknowledgments

RHF and GAK acknowledge grant support from the National Institute on Aging (P30AG067988).

Data Availability

The data used in this study are not openly accessible due to patient privacy, security, and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirement.

Authors' Contributions

Conceptualization: HP, GAK, KL, RHF, XW
Data curation: HP, KL, XW
Formal analysis: HP, KL, LCH, XW
Funding acquisition: GAK, RHF, XW
Methodology: HP, KL, LCH, XW
Writing – original draft: HP, GAK, KL, RHF, XW
Writing – review & editing: HP, GAK, KL, RHF, YSM, XW

Conflicts of Interest

HP, KL, LCH, and XW are current employees of IMO Health. All other authors declare no other competing financial or nonfinancial interests.

Multimedia Appendix 1
ICD (*International Classification of Diseases*) codes used for the phenotyping of dementia cohort.
[DOCX File, 15 KB - aging_v8i1e65221_app1.docx]

Multimedia Appendix 2
Query terms used for identifying memory loss–related symptoms.
[DOCX File, 15 KB - aging_v8i1e65221_app2.docx]

Multimedia Appendix 3
Evaluation of memory loss NLP pipeline. NLP: natural language processing.
[DOCX File, 15 KB - aging_v8i1e65221_app3.docx]

Multimedia Appendix 4
Descriptive statistics of providers and insurance information.
[DOCX File, 14 KB - aging_v8i1e65221_app4.docx]

Multimedia Appendix 5
The distribution time intervals and complaints. (A) Distribution of the time intervals between the first memory loss complaints and the diagnosis of dementia. (B) Distribution of the number of complaints made before the diagnosis of dementia.
[DOCX File, 82 KB - aging_v8i1e65221_app5.docx]

References

1. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology 2013 May 7;80(19):1778-1783. [doi: 10.1212/WNL.0b013e31828726f5] [Medline: 23390181]
2. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6]
3. Matthews KA, Xu W, Gaglioti AH, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. Alzheimers Dement 2019 Jan;15(1):17-24. [doi: 10.1016/j.jalz.2018.06.3063] [Medline: 30243772]
4. Plassman BL, Langa KM, Fisher GG, et al. Prevalence of dementia in the United States: the aging, demographics, and memory study. Neuroepidemiology 2007;29(1-2):125-132. [doi: 10.1159/000109998] [Medline: 17975326]
5. Velandia PP, Miller-Petrie MK, Chen C, et al. Global and regional spending on dementia care from 2000-2019 and expected future health spending scenarios from 2020-2050: an economic modelling exercise. eClinicalMedicine 2022 Mar;45:101337. [doi: 10.1016/j.eclinm.2022.101337] [Medline: 35299657]
6. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):88-106. [doi: 10.1016/S1474-4422(18)30403-4]
7. Frankish H, Horton R. Prevention and management of dementia: a priority for public health. Lancet 2017 Dec;390(10113):2614-2615. [doi: 10.1016/S0140-6736(17)31756-7]
8. Tisher A, Salardini A. A comprehensive update on treatment of dementia. Semin Neurol 2019 Apr;39(2):167-178. [doi: 10.1055/s-0039-1683408] [Medline: 30925610]
9. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell 2019 Oct 3;179(2):312-339. [doi: 10.1016/j.cell.2019.09.001] [Medline: 31564456]
10. Gale SA, Acar D, Daffner KR. Dementia. Am J Med 2018 Oct;131(10):1161-1169. [doi: 10.1016/j.amjmed.2018.01.022] [Medline: 29425707]
11. 2021 Alzheimer's disease facts and figures. Alzheimers Dement 2021 Mar;17(3):327-406. [doi: 10.1002/alz.12328] [Medline: 33756057]
12. Ferencz B, Gerritsen L. Genetics and underlying pathology of dementia. Neuropsychol Rev 2015 Mar;25(1):113-124. [doi: 10.1007/s11065-014-9276-3] [Medline: 25567624]
13. Flicker C, Ferris SH, Reisberg B. Mild cognitive impairment in the elderly: predictors of dementia. Neurology 1991 Jul;41(7):1006-1009. [doi: 10.1212/wnl.41.7.1006] [Medline: 2067629]
14. Ossenkoppele R, Smith R, Mattsson-Carlgren N, et al. Accuracy of tau positron emission tomography as a prognostic marker in preclinical and prodromal Alzheimer disease: a head-to-head comparison against amyloid positron emission tomography and magnetic resonance imaging. JAMA Neurol 2021 Aug 1;78(8):961-971. [doi: 10.1001/jamaneurol.2021.1858] [Medline: 34180956]
15. Vermunt L, Sikkes SAM, van den Hout A, et al. Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. Alzheimers Dement 2019 Jul;15(7):888-898. [doi: 10.1016/j.jalz.2019.04.001] [Medline: 31164314]
16. Amjad H, Roth DL, Sheehan OC, Lyketsos CG, Wolff JL, Samus QM. Underdiagnosis of dementia: an observational study of patterns in diagnosis and awareness in US older adults. J Gen Intern Med 2018 Jul;33(7):1131-1138. [doi: 10.1007/s11606-018-4377-y] [Medline: 29508259]
17. Maclagan LC, Abdalla M, Harris DA, et al. Can patients with dementia be identified in primary care electronic medical records using natural language processing? J Healthc Inform Res 2023 Mar;7(1):42-58. [doi: 10.1007/s41666-023-00125-6] [Medline: 36910911]
18. Cui Y, Liu B, Luo S, et al. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLOS ONE 2011 Jul;6(7):e21896. [doi: 10.1371/journal.pone.0021896]
19. Yao L, Zhang Y, Li Y, Sanseau P, Agarwal P. Electronic health records: implications for drug discovery. Drug Discov Today 2011 Jul;16(13-14):594-599. [doi: 10.1016/j.drudis.2011.05.009] [Medline: 21624499]
20. Ellsworth MA, Dziadzko M, O'Horo JC, Farrell AM, Zhang J, Herasevich V. An appraisal of published usability evaluations of electronic health records via systematic review. J Am Med Inform Assoc 2017 Jan;24(1):218-226. [doi: 10.1093/jamia/ocw046] [Medline: 27107451]
21. Wang Y, Wang L, Rastegar-Mojarad M, et al. Clinical information extraction applications: a literature review. J Biomed Inform 2018 Jan;77:34-49. [doi: 10.1016/j.jbi.2017.11.011] [Medline: 29162496]
22. Bacigalupo I, Lombardo FL, Bargagli AM, et al. Identification of dementia and MCI cases in health information systems: an Italian validation study. Alzheimers Dement 2022;8(1):e12327. [doi: 10.1002/trc2.12327] [Medline: 36320346]
23. Tjandra D, Migrino RQ, Giordani B, Wiens J. Cohort discovery and risk stratification for Alzheimer's disease: an electronic health record-based approach. Alzheimers Dement 2020;6(1):e12035. [doi: 10.1002/trc2.12035] [Medline: 32548236]
24. Fu S, Chen D, He H, et al. Clinical concept extraction: a methodology review. J Biomed Inform 2020 Sep;109:103526. [doi: 10.1016/j.jbi.2020.103526] [Medline: 32768446]

25. Wang X, Chused A, Elhadad N, Friedman C, Markatou M. Automated knowledge acquisition from clinical narrative reports. AMIA Annu Symp Proc 2008 Nov 6;2008:783-787. [Medline: 18999156]
26. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. J Am Med Inform Assoc 2009;16(3):328-337. [doi: 10.1197/jamia.M3028] [Medline: 19261932]
27. Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. J Am Med Inform Assoc 2019 Apr 1;26(4):364-379. [doi: 10.1093/jamia/ocy173] [Medline: 30726935]
28. Lee K, Liu Z, Chandran U, et al. Detecting ground glass opacity features in patients with lung cancer: automated extraction and longitudinal analysis via deep learning–based natural language processing. JMIR AI 2023;2:e44537. [doi: 10.2196/44537] [Medline: 38875565]
29. Weegar R. Applying natural language processing to electronic medical records for estimating healthy life expectancy. Lancet Reg Health West Pac 2021 Apr;9:100132. [doi: 10.1016/j.lanwpc.2021.100132] [Medline: 34327441]
30. Maclagan LC, Abdalla M, Harris DA, et al. Using natural language processing to identify signs and symptoms of dementia and cognitive impairment in primary care electronic medical records (EMR). Alzheimers Dement 2021 Dec;17(S7):e054091. [doi: 10.1002/alz.054091]
31. Al-Harrasi AM, Iqbal E, Tsamakis K, et al. Motor signs in Alzheimer's disease and vascular dementia: detection through natural language processing, co-morbid features and relationship to adverse outcomes. Exp Gerontol 2021 Apr;146:111223. [doi: 10.1016/j.exger.2020.111223] [Medline: 33450346]
32. Byrd RJ, Steinhubl SR, Sun J, Ebadollahi S, Stewart WF. Automatic identification of heart failure diagnostic criteria, using text analysis of clinical notes from electronic health records. Int J Med Inform 2014 Dec;83(12):983-992. [doi: 10.1016/j.ijmedinf.2012.12.005] [Medline: 23317809]
33. Divita G, Luo G, Tran LTT, Workman TE, Gundlapalli AV, Samore MH. General symptom extraction from VA electronic medical notes. Stud Health Technol Inform 2017;245:356-360. [Medline: 29295115]
34. Fu S, Lopes GS, Pagali SR, et al. Ascertainment of delirium status using natural language processing from electronic health records. J Gerontol A Biol Sci Med Sci 2022 Mar 3;77(3):524-530. [doi: 10.1093/gerona/glaa275] [Medline: 35239951]
35. Mahmoudi E, Wu W, Najarian C, Aikens J, Bynum J, Vydiswaran VGV. Identifying caregiver availability using medical notes with rule-based natural language processing: retrospective cohort study. JMIR Aging 2022 Sep 22;5(3):e40241. [doi: 10.2196/40241] [Medline: 35998328]
36. Kharrazi H, Anzaldi LJ, Hernandez L, et al. The value of unstructured electronic health record data in geriatric syndrome case identification. J Am Geriatr Soc 2018 Aug;66(8):1499-1507. [doi: 10.1111/jgs.15411] [Medline: 29972595]
37. Maserejian N, Krzywy H, Eaton S, Galvin JE. Cognitive measures lacking in EHR prior to dementia or Alzheimer's disease diagnosis. Alzheimers Dement 2021 Jul;17(7):1231-1243. [doi: 10.1002/alz.12280] [Medline: 33656251]
38. Noori A, Magdamo C, Liu X, et al. Development and evaluation of a natural language processing annotation tool to facilitate phenotyping of cognitive status in electronic health records: diagnostic study. J Med Internet Res 2022 Aug 30;24(8):e40384. [doi: 10.2196/40384] [Medline: 36040790]
39. Zhou X, Wang Y, Sohn S, Therneau TM, Liu H, Knopman DS. Automatic extraction and assessment of lifestyle exposures for Alzheimer's disease using natural language processing. Int J Med Inform 2019 Oct;130:103943. [doi: 10.1016/j.ijmedinf.2019.08.003] [Medline: 31476655]
40. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-text documentation of dementia symptoms in home healthcare: a natural language processing study. Gerontol Geriatr Med 2020;6. [doi: 10.1177/2333721420959861] [Medline: 33029550]
41. McCoy TH, Han L, Pellegrini AM, Tanzi RE, Berretta S, Perlis RH. Stratifying risk for dementia onset using large-scale electronic health record data: a retrospective cohort study. Alzheimers Dement 2020 Mar;16(3):531-540. [doi: 10.1016/j.jalz.2019.09.084] [Medline: 31859230]
42. McCoy TH Jr, Yu S, Hart KL, et al. High throughput phenotyping for dimensional psychopathology in electronic health records. Biol Psychiatry 2018 Jun 15;83(12):997-1004. [doi: 10.1016/j.biopsych.2018.01.011] [Medline: 29496195]
43. Oh IY, Schindler SE, Ghoshal N, Lai AM, Payne PRO, Gupta A. Extraction of clinical phenotypes for Alzheimer's disease dementia from clinical notes using natural language processing. JAMIA Open 2023 Apr;6(1):ooad014. [doi: 10.1093/jamiaopen/ooad014] [Medline: 36844369]
44. Cheligeer C, Wu G, Lee S, et al. BERT-based neural network for inpatient fall detection from electronic medical records: retrospective cohort study. JMIR Med Inform 2024 Jan 30;12:e48995. [doi: 10.2196/48995] [Medline: 38289643]
45. Ben Miled Z, Dexter PR, Grout RW, Boustani M. Feature engineering from medical notes: a case study of dementia detection. Heliyon 2023 Mar;9(3):e14636. [doi: 10.1016/j.heliyon.2023.e14636] [Medline: 37020943]
46. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on Jan 16, 2013. [doi: 10.48550/ARXIV.1301.3781]
47. Soysal E, Wang J, Jiang M, et al. CLAMP – a toolkit for efficiently building customized clinical natural language processing pipelines. J Am Med Inform Assoc 2018 Mar 1;25(3):331-336. [doi: 10.1093/jamia/ocx132] [Medline: 29186491]

48. Lee K, Liu Z, Chandran U, et al. Detecting ground glass opacity features in patients with lung cancer: automated extraction and longitudinal analysis via deep learning-based natural language processing. JMIR AI 2023 Jun 1;2:e44537. [doi: 10.2196/44537] [Medline: 38875565]
49. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. Alzheimer Dis Assoc Disord 2009;23(4):306-314. [doi: 10.1097/WAD.0b013e3181a6bebc] [Medline: 19568149]
50. Kotagal V, Langa KM, Plassman BL, et al. Factors associated with cognitive evaluations in the United States. Neurology 2015 Jan 6;84(1):64-71. [doi: 10.1212/WNL.0000000000001096] [Medline: 25428689]
51. Harding BN, Floyd JS, Scherrer JF, et al. Methods to identify dementia in the electronic health record: comparing cognitive test scores with dementia algorithms. Healthcare 2020 Jun;8(2):100430. [doi: 10.1016/j.hjdsi.2020.100430] [Medline: 32553526]

Abbreviations

BERT: Bidirectional Encoder Representation from Transformer
BiLSTM: Bidirectional Long Short-Term Memory
CLAMP: Clinical Language Annotation, Modeling, and Processing
CRF: Conditional Random Field
EHR: electronic health record
FN: false negative
FP: false positive
GPCOG: General Practitioner Assessment of Cognition
***ICD*:** *International Classification of Diseases*
MMSE: Mini Mental State Examination
MoCA: Montreal Cognitive Assessment
NLP: natural language processing
OR: odds ratio
PPV: positive predictive value
SLUMS: Saint Louis University Mental Status
TP: true positive

Edited by J Job; submitted 08.08.24; peer-reviewed by K Elechi, M Elbattah, P Han; revised version received 13.01.25; accepted 14.01.25; published 25.02.25.

Please cite as:
Paek H, Fortinsky RH, Lee K, Huang LC, Maghaydah YS, Kuchel GA, Wang X
Real-World Insights Into Dementia Diagnosis Trajectory and Clinical Practice Patterns Unveiled by Natural Language Processing: Development and Usability Study
JMIR Aging 2025;8:e65221
URL: https://aging.jmir.org/2025/1/e65221
doi:10.2196/65221

Factors Influencing the Severity of Urinary and Defecatory Dysfunction Among the Middle-Aged and Older Adult Chinese Population: Longitudinal Study of a 5-Wave Survey Cohort

Haoran Zhu[1], PhD; Xiaoming Li[2], MPH; Liwei Jing[1], PhD; Jingli Kou[3], BSN; Lichao Gong[3], BSN; Rui Wang[1], PhD; Guangtian Liu[4], MSPT; Chao Zhang[5], PhD; Jiayi Zhao[6], MSN; Jing Zhang[7], PhD

[1]School of Nursing, Capital Medical University, 10 Xitoutiao, You'anmenwai, Fengtai District, Beijing, China

[2]School of Public Health, North China University of Science and Technology, Tangshan, China

[3]Department of Geriatrics, Xuanwu Hospital of Capital Medical University, Beijing, China

[4]College of Nursing and Rehabilitation, North China University of Science and Technology, Tangshan, China

[5]Department of Gastroenterology, Peking University Third Hospital, Beijing, China

[6]St Michael College, University of Toronto, Toronto, ON, Canada

[7]School of Public Health, Harbin Medical University, Harbin, China

Corresponding Author:
Liwei Jing, PhD
School of Nursing, Capital Medical University, 10 Xitoutiao, You'anmenwai, Fengtai District, Beijing, China

Abstract

Background: Urinary and defecatory dysfunction (UDD) is a significant concern among the aging population in China. However, there is a lack of longitudinal research exploring the risk factors of UDD severity in Chinese older adults.

Objective: This study uses data from the China Health and Retirement Longitudinal Study spanning 2011 to 2020 to explore UDD risk factors in the middle-aged and older adult Chinese population, focusing on epidemiological characteristics and potential influences on severity.

Methods: A longitudinal cohort of over 10,000 participants from the China Health and Retirement Longitudinal Study was analyzed across 5 waves using Bayesian logistic regression. This analysis examined associations between UDD severity and factors including demographic, lifestyle, and health-related factors, including comorbidities, BMI, and handgrip strength.

Results: Higher UDD prevalence was observed among female population, older adults, those with low education levels, and rural residents. Depression, arthritis, and low handgrip strength emerged as critical predictors of severe UDD. Additionally, abnormal BMI, both underweight (odds ratio [OR] 3.019, 95% CI 1.484 - 5.951; *P*=0.002) and obesity (OR 2.697, 95% CI 1.338 - 5.217; *P*=0.005), was strongly linked to increased severity and persistence of UDD. Participants aged 66 years and older exhibited the highest UDD prevalence, with both underweight and obese individuals facing the greatest risk of persistent and worsening symptoms.

Conclusions: This study is the first to longitudinally examine the risk factors of UDD severity in China's middle-aging and aging population. The findings underscore the need for targeted interventions focusing on muscle strength rehabilitation and comorbidity management to mitigate UDD progression, contributing to improved quality of life for older individuals.

(JMIR Aging 2025;8:e70541) doi:10.2196/70541

KEYWORDS

China; characteristic; risk factors; urination and defecatory dysfunction; longitudinal cohort study

Introduction

Population aging has become a public health issue, imposing a major global challenge across countries in recent decades [1]. The population aged 60 years and older in China is projected to exceed 300 million by 2025, accounting for more than 20% of the total population [2]. Due to the decline in muscle function and disorders of the nervous system among older people, particularly hormonal changes in female population that lead to muscle changes, and urethral sphincter atrophy in male population, millions experience urinary and defecatory dysfunction (UDD) [3-7].

The World Health Organization's integrated care for older people model emphasizes the significance of UDD, particularly urinary incontinence, highlighting its impact on health systems worldwide. Additionally, the care burden associated with urinary retention, fecal incontinence, and constipation is substantial, underscoring the need for comprehensive strategies to address

these pervasive health issues [8]. UDD is defined as a disorder that affects the storage or elimination of urine and feces [9]. Recent studies on UDD reveal varying prevalence rates. Research shows 37.1% global prevalence for female urinary incontinence [10], 18.9% for constipation in older adults [11], and 8% for fecal incontinence [12]. However, these studies, mainly from China, are limited by small, regional samples and a lack of comprehensive data on influencing factors, highlighting the need for broader research to fully understand UDD's impact. This common disorder can have a negative impact on the quality of life of patients and exacerbate their economic burden, particularly in long-term or severe cases [13-16].

Numerous studies have demonstrated that UDD may be associated with chronic conditions (such as hypertension, diabetes, stroke, depression, and so on), abnormal BMI (overweight, obesity, or underweight), handgrip strength, and unhealthy lifestyle factors such as smoking and drinking [10,17-22]. Due to limitations in existing cross-sectional studies, which cannot confirm causal relationships, we established a cohort of over 10,000 participants, with an innovative focus on UDD severity, to study the factors influencing its severity.

The study aimed to use data from a 5-wave survey conducted by the China Health and Retirement Longitudinal Study (CHARLS) [23] to (1) describe the characteristics of UDD from 2011 to 2020 and (2) to identify the risk factors impacting the severity of UDD in China.

Methods

Study Design and Participants

The CHARLS, a nationally representative survey targeting individuals aged 45 years and older from 450 villages or communities in 28 provinces and 150 counties or districts across Mainland China, was conducted by the National School of Development at Peking University from 2011 to 2020 [23]. A total of 150 counties were randomly selected, with stratification by region and urban or rural classification. Administrative villages and neighborhoods were designated as primary sampling units. Households with at least 1 member aged 45 years or older were selected, and 1 individual, along with their spouse, was interviewed. The CHARLS database encompasses a broad array of variables related to demographic statistics, socioeconomic conditions, and health status. Participants in CHARLS are interviewed face-to-face every 2 years using a computer-assisted personal interview technique.

To map the characteristics and temporal trends in UDD prevalence, all participants with UDD were identified in each wave of the cross-sectional survey from 2011 to 2020. For identification of risk factors, participants with UDD who enrolled in 2011 and were continuously followed up in the subsequent 4 cross-sectional surveys were included. Detailed information about the sampling design used in the survey has been described in earlier publications [24].

Ethical Considerations

The CHARLS was approved by the Biomedical Ethics Review Committee of Peking University (IRB0000105211015) [23]. All participants provided signed informed consent forms and received modest financial compensation. Data were anonymized before release and are only available to approved researchers.

Data Collection

In each county or district, trained staff collected data at participants' homes and local community health centers as well as County Centers for Disease Prevention and Control in accordance with the reported protocol [23]. Deidentified information was collected for analysis, consisting of data on participants' demographic characteristics, health-related behaviors and outcomes, childhood conditions, community environment, cognitive and physical health, current economic status, social and family support, health insurance, health care use, comorbidities (chronic diseases such as hypertension and diabetes), and their urinary and defecation control abilities.

Referring to the World Health Organization standard, participants were divided into the following age groups: 45 - 59 years (middle-aged), 60 - 74 years (young-old), 75 - 89 years (old), and ≥90 years (very old) [25]. Residences of participants included rural and urban areas. Education levels were categorized as illiterate, primary school graduate, secondary school graduate, and college graduate or above. Regions of China were classified as Southwest, Southern, Eastern, Northwest, Northern, and Northeast.

The CHARLS did not include data from Ningxia Hui Autonomous Region, Xizang Autonomous Region, Hainan province, Hong Kong, Macau, and Taiwan province. Additionally, the 2020 follow-up survey in the CHARLS did not include participants from the Xinjiang Uygur Autonomous Region due to the impact of the COVID-19 pandemic.

Case Identification and Study Outcomes

According to the International Classification of Functioning, Disability, and Health, urinary function refers to the ability to discharge urine from the bladder, while bowel function refers to the ability to expel waste materials and undigested food in the form of feces, along with the associated physiological processes [26]. Therefore, impairment in either of these functions is classified as UDD. Based on participants' responses to the question "Do you have any difficulties with controlling urination and defecation?" in CHARLS, all participants reporting "I have difficulty" or "I cannot do it" were identified as having UDD.

The second aim of the study was to explore the risk factors for UDD. To this end, participants with UDD were divided into the 4 subgroups based on clinical expert experience and the actual disease status of the participants, each comprising, individuals with increasingly severe UDD, from group A to group D, in order, as follows:

- Group A: Participants who had never experienced UDD.
- Group B: Participants who identified as having UDD in 1 or 2 consecutive surveys, without recurrence in the later follow-ups.
- Group C: Participants who reported having experienced UDD in 1 or 2 consecutive surveys and had recovered at the time of the subsequent 1 or 2 follow-ups but eventually experienced recurrence.

- Group D: Participants who identified as having UDD in consecutive 3 or more surveys.

Statistical Analysis

To achieve the first study aim, a descriptive analysis was conducted to examine the characteristic of prevalence of UDD by sex, residence, marital status, age group, education level, and geographic region over 5 waves from 2011 to 2020. The calculation of the prevalence rate is the cumulative number of cases divided by the total number of participants followed up per year.

To achieve the second study aim, a chi-square test was initially performed to identify potential risk factors among the 4 groups. These factors were then further examined through Bayesian logistic regression analysis. Prior to regression analysis, chi-square tests were conducted to assess univariate associations. A weakly informative prior was used to ensure robust parameter estimation. Parameter estimation was conducted using Markov Chain Monte Carlo sampling, and model results are presented as odds ratios (ORs) with 95% CIs. To enhance interpretability, exponentiation was applied to the estimated coefficients, and 2-sided *P* values were calculated. R (version 64 4.4.3; R Foundation for Statistical Computing) was used for descriptive analysis and logistic regression. *P* values were 2-sided with a significance level of .05.

Results

Characteristics of Study Participants in the 5 Waves

The participants in the 5 waves comprised 17,156, 17,897, 17,715, 19,097, and 19,129 individuals, respectively. Although there were no significant differences between sexes, significant differences were found among age groups, places of residence, education levels, marital status, and regions across the 5 waves. Figure 1 shows the screening process of this study. Table 1 presents the characteristics of patients with UDD across various groups.

Figure 1. Flowchart of participant inclusion or exclusion and grouping diagram. UDD: urinary and defecatory dysfunction.

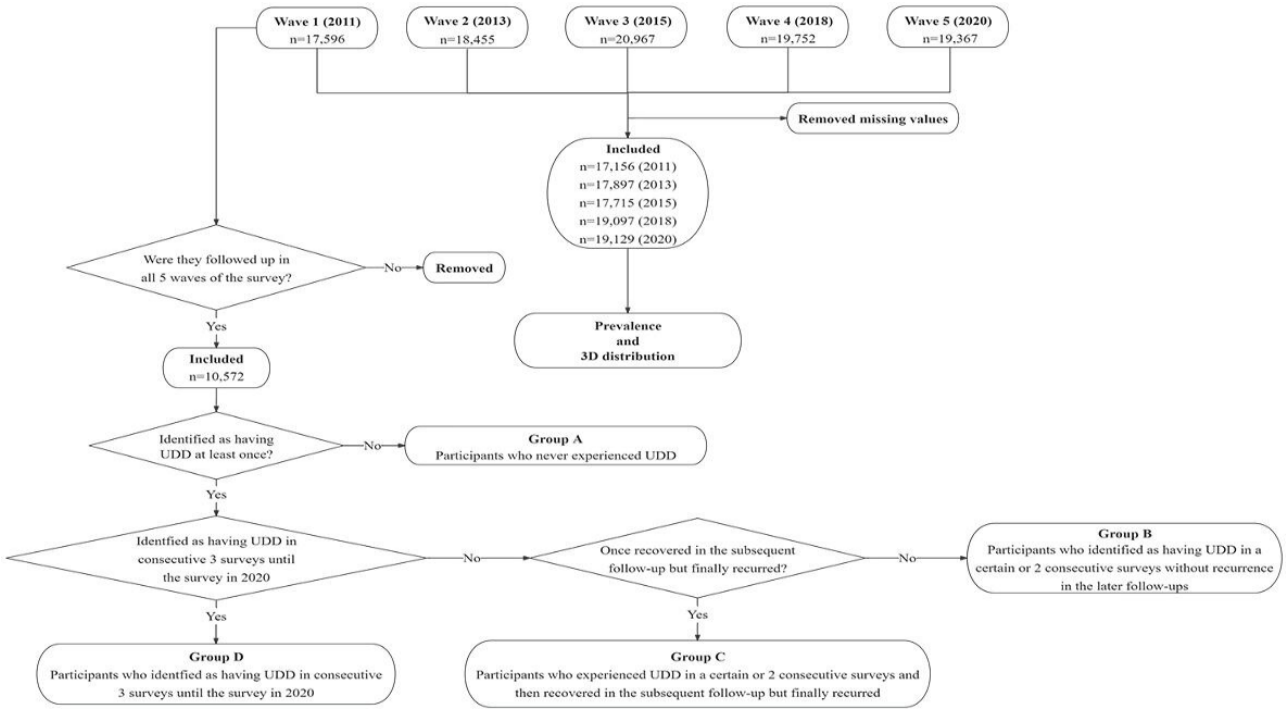


Table . Demographic characteristics of participants with urinary and defecatory dysfunction in the 5-wave survey from 2011 to 2020.

	Baseline wave in 2011 (n=17,156), n (%)	Second wave in 2013 (n=17,897), n (%)	Third wave in 2015 (n=17,715), n (%)	Fourth wave in 2018 (n=19,097), n (%)	Fifth wave in 2020 (n=19,129), n (%)	*P* value
Sex						.98
Female	388 (4.41)	349 (3.77)	429 (4.66)	491 (4.89)	513 (5.09)	
Male	365 (4.63)	372 (4.31)	401 (4.71)	411 (4.54)	447 (4.94)	
Age group (years)						<.001
45 - 50	57 (1.49)	54 (1.46)	47 (1.62)	31 (1.30)	32 (1.27)	
51 - 55	59 (2.11)	58 (2.18)	76 (2.43)	54 (1.47)	73 (1.97)	
56 - 60	130 (3.72)	95 (2.75)	97 (3.34)	67 (2.48)	90 (3.21)	
61 - 65	132 (4.91)	133 (4.31)	136 (4.19)	132 (3.87)	161 (4.69)	
66 - 70	106 (5.92)	104 (5.15)	138 (5.99)	185 (6.24)	199 (6.66)	
71 - 75	92 (7.55)	110 (7.71)	129 (8.40)	134 (7.49)	154 (8.75)	
76 - 80	67 (8.54)	80 (8.71)	111 (11.28)	144 (12.04)	122 (11.37)	
≥81	110 (20.04)	87 (13.30)	96 (13.64)	155 (15.78)	129 (15.64)	
Education						<.001
Illiterate[a]	476 (6.15)	414 (5.09)	482 (5.99)	545 (6.48)	537 (6.51)	
Primary[b]	146 (3.98)	147 (3.88)	177 (4.59)	185 (4.42)	205 (4.90)	
Second[c]	122 (2.29)	150 (2.70)	162 (3)	162 (2.64)	182 (3.13)	
College[d]	9 (2.14)	10 (2.43)	9 (2.17)	10 (2.63)	36 (4.17)	
Marital status						<.001
Married[e]	593 (3.96)	560 (3.61)	625 (4.09)	635 (3.92)	695 (4.34)	
Other status[f]	160 (7.30)	161 (6.78)	205 (8.37)	267 (9.19)	265 (8.51)	
Residence						<.001
Rural	627 (4.72)	585 (4.27)	637 (4.70)	738 (4.87)	728 (5.11)	
City	122 (3.24)	131 (3.27)	180 (4.64)	163 (4.17)	143 (4.99)	
Region						<.001
Northeast	66 (5.18)	52 (3.75)	59 (4.52)	68 (5.16)	64 (5.41)	
East	222 (4.27)	159 (2.93)	201 (3.63)	213 (3.62)	263 (4.38)	
North	87 (3.72)	124 (5.13)	119 (5.19)	146 (5.75)	136 (5.64)	
Central	94 (3.49)	119 (4.26)	126 (4.61)	150 (5.11)	162 (5.32)	
South	42 (2.76)	49 (3.25)	64 (4.37)	58 (3.66)	78 (4.75)	
Southwest	199 (6.84)	163 (5.25)	217 (7.04)	204 (6.06)	185 (5.39)	
Northwest	43 (3.51)	55 (4.40)	44 (3.39)	63 (4.29)	72 (5.18)	

[a]Without formal education, did not finish primary school, was homeschooled.

[b]With primary school education.

[c]With middle school or high school education.

[d]With college education and above.

[e]Married with spouse present or married but not living with spouse temporarily for reasons such as work.

[f]Separated, divorced, widowed, or never married.

Trends and Distributions of UDD Prevalence From 2011 to 2020

The prevalence of UDD from 2011 to 2020 was 4.39% (n=753), 4.03% (n=721), 4.69% (n=830), 4.72% (n=902), and 5.02% (n=960), respectively. Figure 2 demonstrates that the prevalence of UDD by sex and residence rose slowly over the 10 years, with female participants or participants living in rural areas having higher prevalence than their counterparts. The prevalence of UDD by education level and marital status remained essentially stable across the decade in China, although illiterate or married participants had higher prevalence than other categories.

Figure 2. Temporal trends of prevalence of urinary and defecatory dysfunction by sex, education level, marital status, and residence among middle-aged and older people in China from 2011 to 2020.

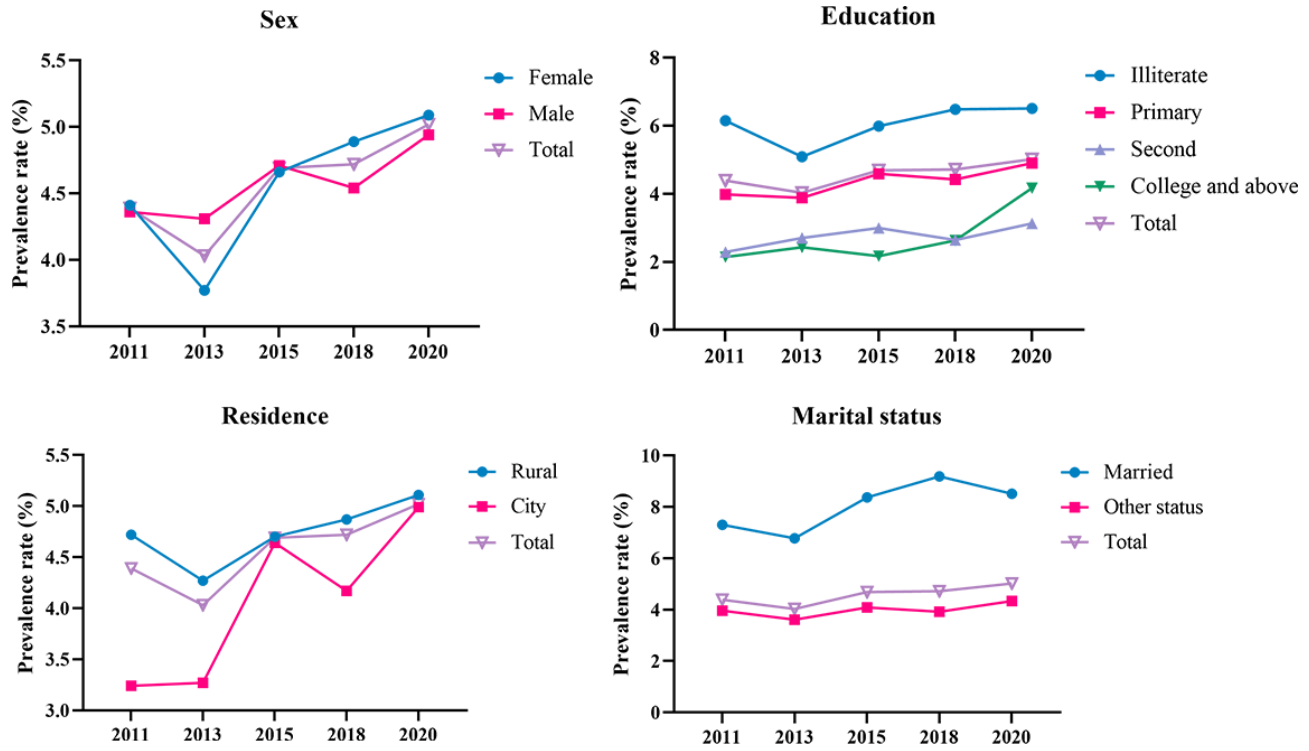
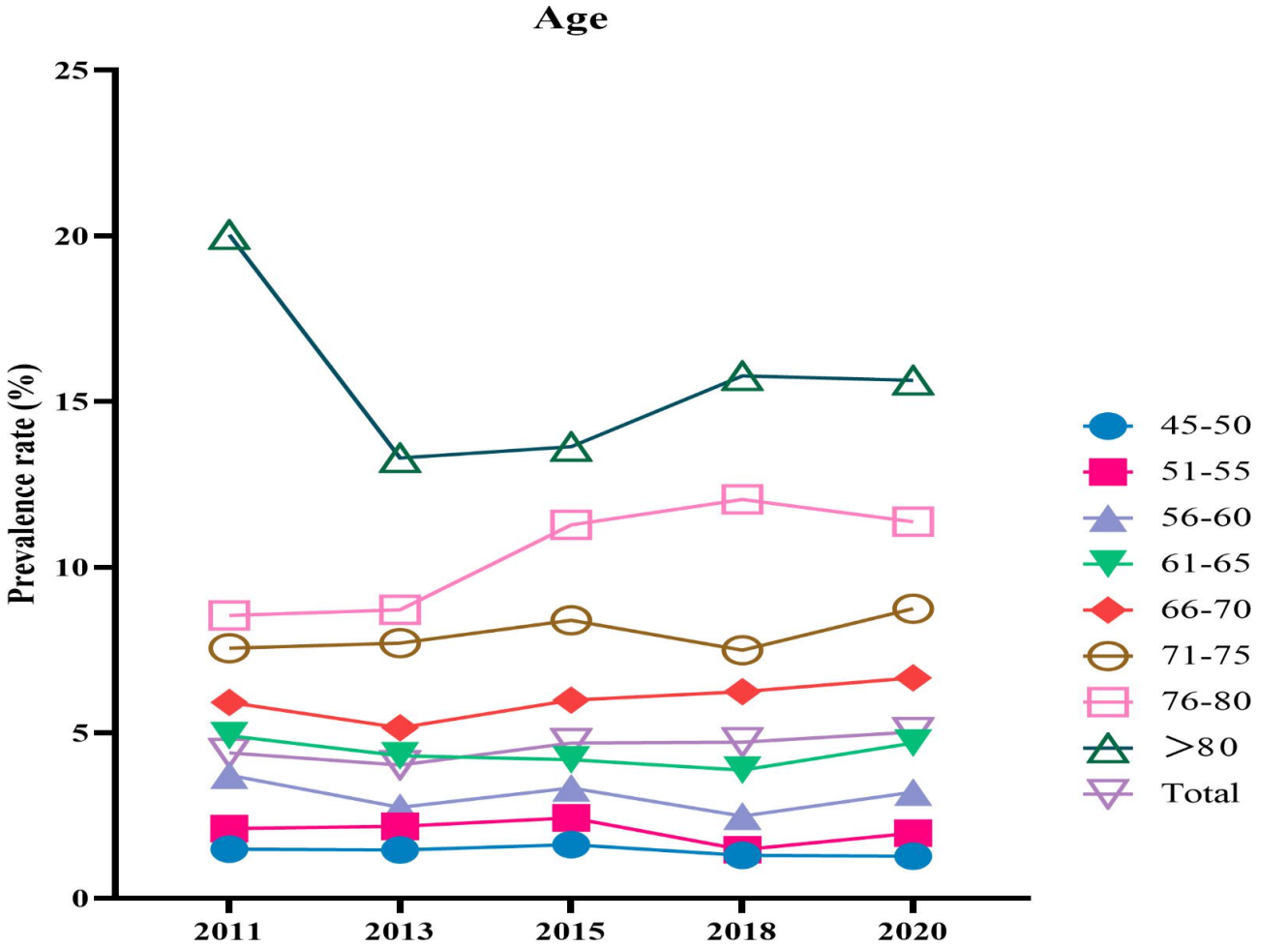


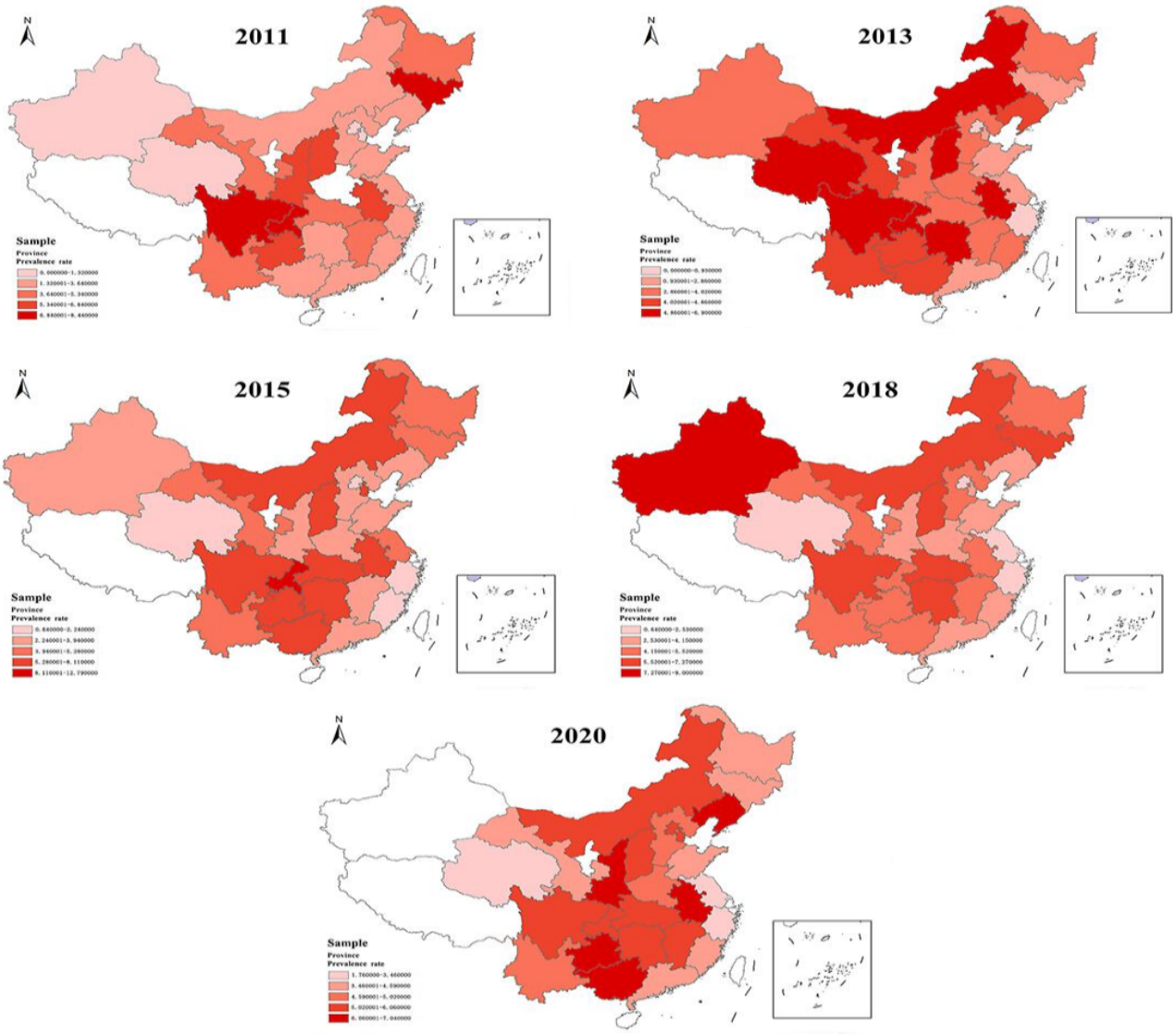
Figure 3 illustrates that the prevalence of UDD gradually increased with age over the decade, particularly among those aged 66 years and older who had a higher prevalence than the overall prevalence. The prevalence of the 4 groups aged 66 years and older slightly increased, while the corresponding figures for the 4 younger groups younger than 65 years were almost stable during the 10 years of surveys.

Figure 3. Temporal trends of age-specific prevalence (in years) of urinary and defecatory dysfunction among middle-aged and older people in China from 2011 to 2020.



The prevalence of UDD varied significantly across different regions over the years, being consistently higher in the Southwest region than in other regions. Both the Southwest and Northwest regions saw an increasing prevalence of UDD in the past 10 years (Figure 4).

Figure 4. Regional distribution of urinary and defecatory dysfunction prevalence among middle-aged and older people in China from 2011 to 2020 (because of the impact of the pandemic, information from Xinjiang participants in 2020 was not included in the China Health and Retirement Longitudinal Study database).



Identification of Risk Factors for Different UDD Conditions in a Cohort Study

There were 10,572 participants enrolled in the cohort study; among them, 8966 participants had never experienced UDD over the 10-year period, while 1275 had recovered, 245 reported recurrence, and 86 reported persistence at the last follow-up. There were significant differences for factors other than sex (P=.19), cancer (P=.12), and smoking (P=.39). Among all groups with various severity of UDD, the proportion of female participants was slightly higher than that of male participants. Higher proportions of participants were aged 60 - 75 years, illiterate, married, living in rural areas, living in the Southwest and East regions, with normal BMI, nonsmokers, and nondrinkers. The comorbidities of participants with UDD are shown in Table 2.

Table . Univariable analysis of 4 groups of urinary and defecatory dysfunction (UDD).

	Without UDD[a] (group A; n=8966), n (%)	Recovered from UDD[b] (group B; n=1275), n (%)	Reported recurrent[c] UDD (group C; n=245), n (%)	Reported persistent UDD[d] (group D; n=86), n (%)	*P* value[e]
Sex					
Female	4751 (52.99)	717 (56.24)	132 (53.88)	47 (54.65）	.19
Age group (years)					<.001
45 - 59	5604 (62.50)	555 (43.53)	85 (34.69)	30 (34.88)	
60 - 75	3015 (33.63)	608 (47.69)	137 (55.92)	47 (54.65)	
76 - 90	343 (3.83)	109 (8.55)	22 (8.98)	8 (9.30)	
＞90	4 (0.04)	3 (0.24)	1 (0.41)	1 (1.16)	
Education					<.001
Illiterate	3935 (43.89)	715 (56.08)	147 (60)	47 (54.65)	
Primary	1967 (21.94)	289 (22.67)	46 (18.78)	26 (30.23)	
Second	2918 (32.55)	260 (20.39)	50 (20.41)	12 (13.95)	
College and above	146 (1.63)	11 (0.86)	2 (0.82)	1 (1.16)	
Marital status					
Married	8169 (91.11)	1088 (85.33)	210 (85.71)	74 (86.05)	<.001
Residence					
Rural	7400 (82.53)	1115 (87.45)	213 (86.94)	74 (86.05)	.002
Region					<.001
Northeast	557 (6.21)	75 (5.88)	15 (6.12)	6 (6.98)	
East	2891 (32.24)	351 (27.53)	63 (25.71)	22 (25.58)	
North	1149 (12.82)	183 (14.35)	30 (12.24)	11 (12.79)	
Central	1482 (16.53)	215 (16.86)	40 (16.33)	14 (16.28)	
South	735 (8.20)	84 (6.59)	13 (5.31)	4 (4.65)	
Southwest	1521 (16.96)	279 (21.88)	68 (27.76)	26 (30.23)	
Northwest	631 (7.04)	88 (6.90)	16 (6.53)	3 (3.49)	
BMI[e]					<.001
Underweight	466 (5.41)	83 (6.83)	24 (10.57)	14 (17.07)	
Normal weight	4486 (52.09)	636 (52.35)	113 (49.78)	29 (35.37)	
Overweight	2615 (30.36)	330 (27.16)	64 (28.19)	20 (24.39)	
Obese	1018 (11.82)	157 (12.92)	25 (11.01)	17 (20.73)	
Smoking[e]	3434 (38.32)	483 (37.97)	107 (43.67)	33 (38.37)	.39
Drinking[e]	3065 (34.21)	342 (26.89)	83 (33.88)	25 (29.07)	<.001
Comorbidities					
Hypertension[e]	1882 (21.09)	387 (30.57)	71 (29.22)	32 (37.21)	<.001
Dyslipidemia[e]	772 (8.64)	144 (11.31)	30 (12.35)	11 (12.79)	.002
Diabetes[e]	424 (4.73)	83 (6.52)	22 (9.05)	10 (11.63)	<.001
Cancer	76 (0.85)	19 (1.49)	1 (0.41)	1 (1.16)	.12
Chronic lung disease[e]	701 (7.84)	184 (14.51)	40 (16.39)	20 (23.26)	<.001
Liver disease[e]	320 (3.57)	59 (4.63)	16 (6.53)	3 (3.49)	.03

	Without UDD[a] (group A; n=8966), n (%)	Recovered from UDD[b] (group B; n=1275), n (%)	Reported recurrent[c] UDD (group C; n=245), n (%)	Reported persistent UDD[d] (group D; n=86), n (%)	*P* value[e]
Heart problems[e]	854 (9.53)	210 (16.51)	45 (18.44)	18 (21.18)	<.001
Stroke[e]	125 (1.40)	41 (3.24)	12 (4.92)	6 (7.06)	<.001
Kidney disease[e]	472 (5.27)	110 (8.63)	35 (14.29)	14 (16.28)	<.001
Digestive disease[e]	1954 (21.86)	370 (29.20)	96 (39.18)	37 (43.53)	<.001
Memory-related disease[e]	66 (0.74)	26 (2.05)	12 (4.90)	6 (6.98)	<.001
Arthritis or rheumatism[e]	2839 (31.74)	578 (45.62)	141 (57.55)	50 (58.14)	<.001
Asthma[e]	239 (2.68)	53 (4.18)	24 (9.84)	6 (6.98)	<.001
Depression[e]	3334 (37.18)	748 (58.67)	174 (71.02)	65 (75.58)	<.001
Handgrip strength					
Low handgrip strength	1005 (11.84)	289 (24.41)	70 (31.39)	30 (37.04)	<.001

[a]Participants who had never experienced UDD.

[b]Participants who identified as having UDD in 1 or 2 consecutive surveys without recurrence in the later follow-ups.

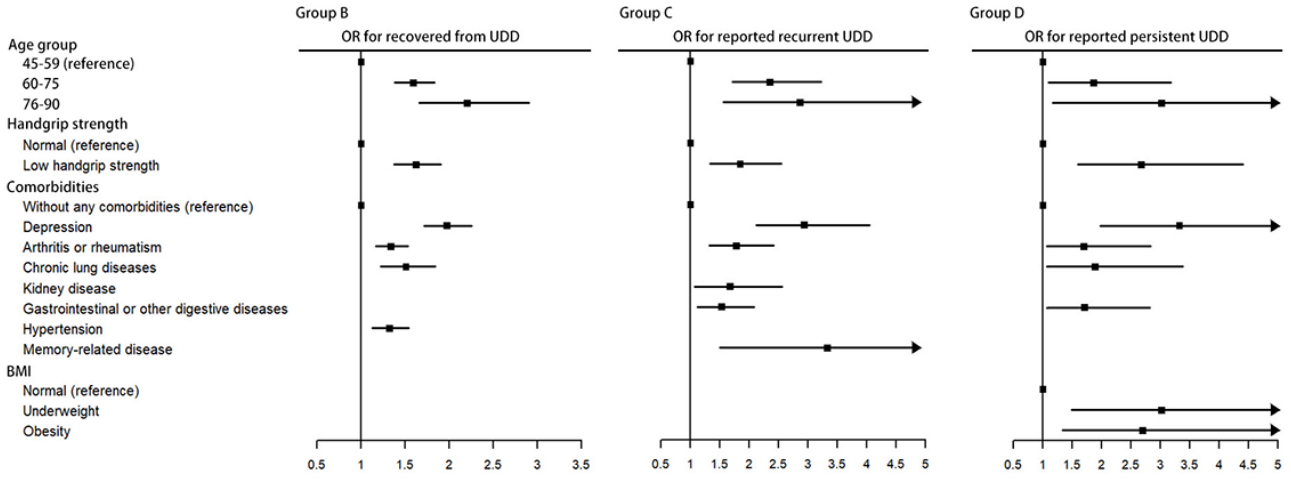
[c]Participants who reported experiencing UDD in 1 or 2 consecutive surveys and having recovered in the subsequent 1 or 2 follow-ups but eventually experienced recurrence.

[d]Participants who identified as having UDD in 3 or more consecutive surveys.

[e]Missing data: 436 for BMI, 7 for smoking, 10 for drinking, 53 for hypertension, 30 for dyslipidemia, 14 for diabetes, 27 for chronic lung disease, 4 for liver disease, 13 for heart problem, 36 for stroke, 5 for kidney disease, 37 for digestive disease, 42 for memory-related disease, 29 for arthritis or rheumatism, 52 for asthma, and 593 for handgrip strength.

Figure 5 shows the ORs and 95% CIs of the risk factors for UDD in patients with different conditions. Four common risk factors were identified across the 3 groups, namely, older age (60 - 89 years), depression, arthritis (or rheumatism), and handgrip strength. For group C and group D, who had more severe UDD than group B, participants with gastrointestinal diseases had a higher risk of experiencing recurrent or persistent UDD. Memory-related diseases (OR 3.328, 95% CI 1.505 - 6.836; *P*=.002) may cause UDD to recur. Additionally, underweight (OR 3.019, 95% CI 1.484 - 5.951; *P*=.002) and obesity (OR 2.697, 95% CI 1.338 - 5.217; *P*=.005) were identified as potential independent risk factors for the persistent nature of UDD among middle-aged and older people in China. The detailed results of the Bayesian logistic regression model are provided in Multimedia Appendix 1.

Figure 5. OR and 95% CI of risk factors for UDD in different population groups with UDD. Forest plots show ORs and 95% CIs for group (B) recovered from UDD, (C) reported UDD recurrence, and (D) UDD persisted adjusted for sex, age, education, marital status, residence, region, BMI, handgrip strength, smoking status, drinking status, and comorbidities. In the forest plot, only the significant results within each group (*P*<.05) are shown. The reference group for groups B, C, and D consists of participants who had never experienced UDD. OR: odds ratio; UDD: urinary and defecatory dysfunction.



Discussion

Principal Findings

This is the first study to use a representative sample to describe the temporal, spatial, and interpersonal distributions of UDD among middle-aged and older people in China. The study identified that over the past decade, the prevalence of UDD has remained stable across different education levels, kinds of marital status, and age groups. However, it significantly increased among both sexes, urban residents, and in the Northwest and Southwest regions. Participants who were female, illiterate, or married and those aged 66 years and older exhibited a higher prevalence of UDD. The incidence, recurrence, and persistence of UDD were more likely to be influenced by aging and comorbidities such as depression, arthritis (or rheumatic diseases), and handgrip strength in China. More importantly, being underweight or obese might contribute to the long-term persistence of UDD.

The Prevalence and Time Trends of UDD

Due to its increased prevalence, the disease burden associated with UDD has shown a gradually escalating trend in China over the past 10 years. However, despite a slight increase, the prevalence of UDD in China (5.02% in 2020) has remained lower than the global prevalence, which ranges from 7% to 55%. This lower prevalence might be attributed to lower self-report due to participants' potential embarrassment or their belief that UDD is a natural aging-related phenomenon among the Chinese population [10,21,27,28]. Approximately, 50% to 67% of patients were unwilling to report their UDD conditions to health care providers [29,30]. The slowly increasing trend in prevalence observed in this study may reflect improved awareness and diagnosis of UDD over time, leading to more UDD cases being identified.

Over the past decade, the prevalence of UDD in female individuals has consistently remained higher than that in male individuals. This finding underscores the necessity of addressing the unique physiological and reproductive health challenges that female individuals face, as the relaxation of pelvic floor muscles during childbirth significantly contributes to this increased prevalence [31]. This suggests that we should place greater emphasis on female individual's pelvic health. By implementing early interventions and training, such as pelvic floor muscle exercises, we can significantly improve the quality of life for female individuals in their later years [32].

It was notable that the Northwest and Southwest provinces—2 resource-limited regions in China—experienced an obviously higher prevalence of UDD during the 5-wave survey. The evidence aligned with other studies demonstrating that economic status may influence the prevalence of UDD, while UDD places a significant economic burden on patients too [14,15,33]. The higher prevalence of UDD among participants from resource-limited areas may be related to the insufficient health insurance or poor access to medical services, such as postpartum pelvic floor rehabilitation [34,35]. Therefore, economic development may serve as a solution to improve the quality of life of patients with UDD in these regions along with enhancing the accessibility of medical resources.

Given the impact of UDD, it is necessary to establish a surveillance system for UDD to provide more evidence for the identification of the long-term impact of UDD and the development of prevention and treatment strategies at the population level.

Population Aging, Comorbidities, and Low Handgrip Strength May Lead to All Severity of UDD Among the Chinese Population as Marked Risk Factors

Population aging is an important risk factor influencing the occurrence, recurrence, and persistence of UDD. This study suggests that the physiological senescence associated with aging may contribute to the development and exacerbation of UDD, particularly among participants aged 66-90 years. It is widely recognized that aging can lead to the occurrence or aggravation of UDD, likely due to the frailty and functional impairment associated with aging [13,36-39]. However, the potential mechanism underlying the association between frailty and UDD remains unknown. Finite element models can make it possible to explore the mechanism underpinning the relationship between frailty and UDD [40]. By comparing defecation outcomes under different parameter settings, the association between training methods and urinary and fecal control ability can be quantified, providing accurate guidance for rehabilitation.

This study demonstrated that comorbidities are the second important risk factor for UDD. For instance, depression was found to significantly impact the incidence, recurrence, and persistence of UDD, as it was usually associated with reduced serotonin function, which leads to urgency urinary incontinence [41]. Some studies argue that the medicines patients take potentially result in gastrointestinal side effects such as constipation [42]. In addition, participants with urinary and fecal incontinence and constipation exhibit significantly higher levels of depression and stress [43-49], possibly due to the production of certain gut microbiota that lead to depressive symptoms and lower levels of *g_Pseudoramibacter-Eubacterium* and *g_Candidatus-Solibacter*. [50] Functional constipation in middle-aged and older individuals may lead to a decrease in the abundance of these microbiota [51], ultimately exacerbating depression. This interaction creates a vicious cycle between depression and UDD.

In addition to depression, arthritis or rheumatism likely also contributes to the increased risk of incidence, recurrence, and persistence of UDD, although its impact on UDD might have been underestimated over the past 20 years. The most recent studies on the association between arthritis and UDD were conducted in the 1990s. These studies implied that arthritis may increase susceptibility to urinary tract infections, and atlantoaxial subluxation in the late stage of arthritis may lead to neurogenic bladder [52], which in turn increases the incidence of UDD. Medications such as misoprostol and cyclophosphamide, used to treat arthritis, may also cause UDD [52,53]. Moreover, opioids, commonly taken by patients with arthritis and rheumatic diseases, are likely to cause constipation. Additionally, the limited mobility of patients with osteoarthritis is perhaps associated with urinary incontinence [54,55].

Other comorbidities, such as hypertension and memory-related diseases, were also identified as potential contributors to the occurrence or recurrence of UDD. For instance, studies indicate a correlation between hypertension and constipation, potentially linked by shared physiological mechanisms and lifestyle factors [56]. Patients with Alzheimer disease often lose the ability to send bladder signals to the brain's urination center, resulting in an inability to urinate normally [57,58].

This study, therefore, suggested potential causal relationships between comorbidities and UDD, indicating that health care providers should pay more attention to the prevention and treatment of these comorbidities. Optimizing pharmaceutical therapy may effectively alleviate the incidence and progression of UDD in middle-aged and older individuals. Additionally, increasing physical therapy or dietary and nutritional treatments for mental health, arthritis or rheumatic diseases, and other chronic conditions could be beneficial. Dissemination of knowledge about these chronic diseases to enhance public awareness at the community level might help address the occurrence, recurrence, and persistence of UDD.

Our study also reveals a significant correlation between grip strength and the severity of UDD. Specifically, individuals with reduced grip strength exhibited a higher incidence of severe UDD manifestations, including recurrent and persistent conditions. This relationship may suggest that grip strength serves as an indirect biomarker for overall muscle integrity and functional capacity, which could influence the severity of UDD symptoms [59,60]. These results align with existing literature emphasizing the importance of muscle function in maintaining urogenital health and also underscore the need for further research into preventive strategies focusing on strength training in vulnerable populations.

Being Underweight or Obese Can Lead to Long-Term Persistence of UDD

This study found that almost half of the older Chinese population in the cohort had abnormal weight, with underweight or obesity considered independent risk factors for persistent UDD. This finding aligned with evidence from other studies showing that obesity is closely related to the occurrence of UDD [10,61-63]. The mechanism underlying this association is that metabolic changes and increased abdominal pressure caused by obesity may lead to the development of UDD. Additionally, underweight older adults might experience reduced muscle strength, particularly in the pelvic floor, possibly triggering the incidence of UDD. Some studies argue that population aging–related muscle atrophy and decreased strength are more likely to cause persistence of UDD [64,65], particularly among those who are underweight. Consequently, the improvement of nutritional intake is vital for older people in both communities and health care facilities. The awareness of healthy diet and appropriate physical exercises be enhanced among the older population, and health care providers and community health workers should develop individual tailored interventions for patients with UDD who are underweight or obese.

Strengths and Limitations

The study population was selected from 28 provinces in China, ensuring sufficient representativeness to reflect the characteristics and risk factors of UDD in the middle-aged and older Chinese populations. To achieve the second study objective, a cohort study was designed to establish clear causal relationships between the identified risk factors and the incidence, progression, and persistence of UDD in the Chinese population. Additionally, the population with UDD was classified based on the severity of UDD, facilitating the exploration of risk factors and the development of personalized interventions.

There are some limitations to this study. The CHARLS does not provide detailed information on urinary and fecal incontinence or constipation, making it difficult to distinguish risk factors for diverse UDD conditions. The CHARLS also lacks information on factors that may affect UDD, such as bladder outlet obstruction, urinary tract infections, benign prostatic hyperplasia, use of analgesics, and history of abuse, and factors such as hospital capacity, network scale, and accessibility of medical services over the past decade could not be analyzed due to the same limitations. Additionally, due to the statistical methods used in the CHARLS database, the impact of incontinence related to childbirth in female population may have been underestimated. Consequently, this study could not assess the impact of these factors on UDD. Furthermore, those who had UDD in both 2018 and 2020 could not have their UDD prognosis determined after the CHARLS survey: they may have recovered and not be included in the most-severe group, which may have resulted in misclassification bias. In addition, the emergence of recovered participants may also be related to respondents' misjudgment of their own UDD conditions. In the Results section, middle school and high school education emerged as a protective factor for the recovered population, though the underlying reasons require further investigation.

Conclusions

In summary, the prevalence of UDD increased with age and was found to be higher among illiterate individuals, married people, and those living in the Southwest and Northwest regions. Depression, arthritis and rheumatic diseases, and other chronic comorbidities contributed to the occurrence, recurrence, and persistence of UDD; further, being underweight or obese independently affected the persistence of UDD among the middle-aged and older Chinese population. Enhancing the treatment of psychological and chronic diseases and improving BMI may alleviate the occurrence, recurrence, and persistence of UDD.

Acknowledgments

This work was supported by the National Key R&D Program of China under the project "Development of Enhanced Training and Nursing Care Robot Systems for Bedridden Elderly with Urinary and Defecatory Dysfunction" (2022YFB4703300). The

authors would like to thank Fan Xuanxuan as well as Xu Ruoyu, Yang Muqi, and Yuan Qi from the School of Nursing, Capital Medical University, for their assistance in data processing and analysis (listed in no particular order).

Data Availability

The data used in this study are publicly available from the China Health and Retirement Longitudinal Study.

Authors' Contributions

HZ designed the study and drafted the manuscript. LJ supervised the study and approved the final version. J Zhang interpreted the data and reviewed the manuscript. GL and XL helped analyze the data. JK contributed to the conceptual framework and provided resources. LG administered the project and funding. RW contributed to methodological expertise and data curation. J Zhao and CZ validated the results. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The detailed results of the Bayesian logistic regression model.
[DOCX File, 31 KB - aging_v8i1e70541_app1.docx]

References

1. Ferrucci L, Gonzalez-Freire M, Fabbri E, et al. Measuring biological aging in humans: a quest. Aging Cell 2020 Feb;19(2):e13080. [doi: 10.1111/acel.13080] [Medline: 31833194]
2. Wang H. Prioritizing the health of the elderly population in China to promote the national strategy for healthy aging. Chin J Epidemiol 2019;40:259-265.
3. Ness W. Faecal incontinence: causes, assessment and management. Nurs Stand 2012;26(42):52-54. [doi: 10.7748/ns2012.06.26.42.52.c9162] [Medline: 22908765]
4. Coyne KS, Sexton CC, Bell JA, et al. The prevalence of lower urinary tract symptoms (LUTS) and overactive bladder (OAB) by racial/ethnic group and age: results from OAB-POLL. Neurourol Urodyn 2013 Mar;32(3):230-237. [doi: 10.1002/nau.22295] [Medline: 22847394]
5. Goepel M, Hoffmann JA, Piro M, Rübben H, Michel MC. Prevalence and physician awareness of symptoms of urinary bladder dysfunction. Eur Urol 2002 Mar;41(3):234-239. [doi: 10.1016/s0302-2838(02)00014-3] [Medline: 12180221]
6. Markland AD, Goode PS, Redden DT, Borrud LG, Burgio KL. Prevalence of urinary incontinence in men: results from the national health and nutrition examination survey. J Urol 2010 Sep;184(3):1022-1027. [doi: 10.1016/j.juro.2010.05.025] [Medline: 20643440]
7. Whitehead WE, Borrud L, Goode PS, et al. Fecal incontinence in US adults: epidemiology and risk factors. Gastroenterology 2009 Aug;137(2):512-517. [doi: 10.1053/j.gastro.2009.04.054] [Medline: 19410574]
8. Integrated care for older people: guidelines on community-level interventions to manage declines in intrinsic capacity. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/9789241550109 [accessed 2025-05-15]
9. Elder JS, Diaz M. Vesicoureteral reflux—the role of bladder and bowel dysfunction. Nat Rev Urol 2013 Nov;10(11):640-648. [doi: 10.1038/nrurol.2013.221] [Medline: 24126731]
10. Batmani S, Jalali R, Mohammadi M, Bokaee S. Prevalence and factors related to urinary incontinence in older adults women worldwide: a comprehensive systematic review and meta-analysis of observational studies. BMC Geriatr 2021 Mar 29;21(1):212. [doi: 10.1186/s12877-021-02135-8] [Medline: 33781236]
11. Salari N, Ghasemianrad M, Ammari-Allahyari M, Rasoulpoor S, Shohaimi S, Mohammadi M. Global prevalence of constipation in older adults: a systematic review and meta-analysis. Wien Klin Wochenschr 2023 Aug;135(15-16):389-398. [doi: 10.1007/s00508-023-02156-w] [Medline: 36826591]
12. Mack I, Hahn H, Gödel C, Enck P, Bharucha AE. Global prevalence of fecal incontinence in community-dwelling adults: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2024 Apr;22(4):712-731. [doi: 10.1016/j.cgh.2023.09.004] [Medline: 37734583]
13. Vaughan CP, Markland AD. Urinary incontinence in women. Ann Intern Med 2020 Feb 4;172(3):ITC17-ITC32. [doi: 10.7326/AITC202002040] [Medline: 32016335]
14. Milsom I, Coyne KS, Nicholson S, Kvasz M, Chen CI, Wein AJ. Global prevalence and economic burden of urgency urinary incontinence: a systematic review. Eur Urol 2014 Jan;65(1):79-95. [doi: 10.1016/j.eururo.2013.08.031] [Medline: 24007713]
15. Miner PB Jr. Economic and personal impact of fecal and urinary incontinence. Gastroenterology 2004 Jan;126(1 Suppl 1):S8-13. [doi: 10.1053/j.gastro.2003.10.056] [Medline: 14978633]

16. Frigerio M, Barba M, Cola A, et al. Quality of life, psychological wellbeing, and sexuality in women with urinary incontinence—where are we now: a narrative review. Med Bogota Colomb 2022;58(4):525. [doi: 10.3390/medicina58040525]
17. Paquette IM, Varma MG, Kaiser AM, Steele SR, Rafferty JF. The American Society of Colon and Rectal Surgeons' clinical practice guideline for the treatment of fecal incontinence. Dis Colon Rectum 2015;58(7):623-636. [doi: 10.1097/DCR.0000000000000397]
18. Drake MJ, Williams J, Bijos DA. Voiding dysfunction due to detrusor underactivity: an overview. Nat Rev Urol 2014 Aug;11(8):454-464. [doi: 10.1038/nrurol.2014.156] [Medline: 25002201]
19. Chaichanavichkij P, Hartmann M, Scott SM, Fenton N, Knowles CH, Disorders of Defaecation Delphi Group. Evaluating the risk factors for the development of benign disorders of defaecation: a surgical perspective. Tech Coloproctol 2023 Oct;27(10):847-857. [doi: 10.1007/s10151-023-02843-w] [Medline: 37498418]
20. Heitmann PT, Vollebregt PF, Knowles CH, Lunniss PJ, Dinning PG, Scott SM. Understanding the physiology of human defaecation and disorders of continence and evacuation. Nat Rev Gastroenterol Hepatol 2021 Nov;18(11):751-769. [doi: 10.1038/s41575-021-00487-5] [Medline: 34373626]
21. Bag Soytas R, Soytas M, Danacioglu YO, et al. Relationship between the types of urinary incontinence, handgrip strength, and pelvic floor muscle strength in adult women. Neurourol Urodyn 2021 Aug;40(6):1532-1538. [doi: 10.1002/nau.24699] [Medline: 34004030]
22. Barberio B, Judge C, Savarino EV, Ford AC. Global prevalence of functional constipation according to the Rome criteria: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2021 Aug;6(8):638-648. [doi: 10.1016/S2468-1253(21)00111-4] [Medline: 34090581]
23. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb;43(1):61-68. [doi: 10.1093/ije/dys203] [Medline: 23243115]
24. Zhao Y, Strauss J, Yang G, et al. 2011–2012 national baseline users' guide. China Health and Retirement Longitudinal Study. 2013. URL: https://charls.pku.edu.cn/wd/a2011nqgjxdc.htm [accessed 2025-05-15]
25. Ageing and health. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-06-18]
26. International Classification of Functioning, Disability and Health: ICF. World Health Organization. 2001. URL: https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [accessed 2024-06-24]
27. Mahfouz IA, Blanker MH, Asali F, et al. Seeking consultation for urinary incontinence: behaviours and barriers among Jordanian women. Neurourol Urodyn 2023 Aug;42(6):1299-1310. [doi: 10.1002/nau.25189] [Medline: 37186396]
28. Jarbøl DE, Haastrup PF, Rasmussen S, Søndergaaard J, Balasubramaniam K. Women's barriers for contacting their general practitioner when bothered by urinary incontinence: a population-based cross-sectional study. BMC Urol 2021 Jul 12;21(1):99. [doi: 10.1186/s12894-021-00864-x] [Medline: 34247613]
29. Walden AL, Salsbury SA, Reed WR, Lawrence DJ. Bladder and bowel symptoms among adults presenting with low back pain to an academic chiropractic clinic: results of a preliminary study. J Chiropr Med 2014 Sep;13(3):178-187. [doi: 10.1016/j.jcm.2014.07.006] [Medline: 25225466]
30. Yan F, Xiao LD, Zhou K, Li Z, Tang S. Perceptions and help-seeking behaviours among community-dwelling older people with urinary incontinence: a systematic integrative review. J Adv Nurs 2022 Jun;78(6):1574-1587. [doi: 10.1111/jan.15183] [Medline: 35150161]
31. Sheng Y, Carpenter JS, Ashton-Miller JA, Miller JM. Mechanisms of pelvic floor muscle training for managing urinary incontinence in women: a scoping review. BMC Womens Health 2022 May 13;22(1):161. [doi: 10.1186/s12905-022-01742-w] [Medline: 35562699]
32. Curillo-Aguirre CA, Gea-Izquierdo E. Effectiveness of pelvic floor muscle training on quality of life in women with urinary incontinence: a systematic review and meta-analysis. Medicina (Kaunas) 2023 May 23;59(6):1004. [doi: 10.3390/medicina59061004] [Medline: 37374208]
33. Harris LA, Chang CH. Burden of constipation: looking beyond bowel movements. Am J Gastroenterol 2022 Apr 1;117(4S):S2-S5. [doi: 10.14309/ajg.0000000000001708] [Medline: 35354769]
34. Barakat C, Konstantinidis T. A review of the relationship between socioeconomic status change and health. Int J Environ Res Public Health 2023 Jun 29;20(13):6249. [doi: 10.3390/ijerph20136249] [Medline: 37444097]
35. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020;8:231. [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
36. Tamanini JTN, Franceschi Júnior O, Santos JLF, et al. Fecal incontinence: incidence and risk factors from the SABE (Health, Wellbeing and Aging) study. Int Urogynecol J 2022 Nov;33(11):2993-3004. [doi: 10.1007/s00192-021-04914-8] [Medline: 35015091]
37. Bauer SR, McCulloch CE, Cawthon PM, et al. Longitudinal associations between concurrent changes in phenotypic frailty and lower urinary tract symptoms among older men. The Journal of Frailty & Aging 2023 Apr;12(2):117-125. [doi: 10.14283/jfa.2022.33]
38. Liu X, Zhao D, Zhao S, Li Z, Wang Y, Qin X. Deciphering the correlations between aging and constipation by metabolomics and network pharmacology. Aging (Albany NY) 2021 Jan 10;13(3):3798-3818. [doi: 10.18632/aging.202340] [Medline: 33428599]

39. Veronese N, Soysal P, Stubbs B, et al. Association between urinary incontinence and frailty: a systematic review and meta-analysis. Eur Geriatr Med 2018 Oct;9(5):571-578. [doi: 10.1007/s41999-018-0102-y] [Medline: 34654231]
40. Wang R, Liu G, Jing L, Zhang J, Ye Y, Zhu H. Quantifying the effects of five rehabilitation training methods on the ability of elderly men to control bowel movements: a finite element analysis study. Front Bioeng Biotechnol 2024;12:1392448. [doi: 10.3389/fbioe.2024.1392448] [Medline: 38988865]
41. Zorn BH, Montgomery H, Pieper K, Gray M, Steers WD. Urinary incontinence and depression. J Urol 1999 Jul;162(1):82-84. [doi: 10.1097/00005392-199907000-00020] [Medline: 10379745]
42. Oliva V, Lippi M, Paci R, et al. Gastrointestinal side effects associated with antidepressant treatments in patients with major depressive disorder: a systematic review and meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry 2021 Jul 13;109:110266. [doi: 10.1016/j.pnpbp.2021.110266] [Medline: 33549697]
43. Lee HY, Rhee Y, Choi KS. Urinary incontinence and the association with depression, stress, and self-esteem in older Korean Women. Sci Rep 2021 Apr 27;11(1):9054. [doi: 10.1038/s41598-021-88740-4] [Medline: 33907278]
44. Wu S, Wu F. Association of urinary incontinence with depression among men: a cross-sectional study. BMC Public Health 2023;23(1):944. [doi: 10.1186/s12889-023-15961-9]
45. Wang Y, Li N, Zhou Q, Wang P. Fecal incontinence was associated with depression of any severity: insights from a large cross-sectional study. Int J Colorectal Dis 2023 Nov 22;38(1):271. [doi: 10.1007/s00384-023-04563-x] [Medline: 37991579]
46. Gevelinger MM, Jaworski EM, Speranza JR, Sanderson DJ. Improvements in self-reported depression following treatment of fecal incontinence with sacral neuromodulation. Neuromodulation 2020 Dec;23(8):1158-1163. [doi: 10.1111/ner.13249]
47. Whiting D, Shane AI, Pope R, Payne S, Venn S. Female urinary incontinence in sub-Saharan Africa. BJU Int 2022 Nov;130(5):543-549. [doi: 10.1111/bju.15903] [Medline: 36161452]
48. Ballou S, Katon J, Singh P, et al. Chronic diarrhea and constipation are more common in depressed individuals. Clin Gastroenterol Hepatol 2019 Dec;17(13):2696-2703. [doi: 10.1016/j.cgh.2019.03.046] [Medline: 30954714]
49. Wang Y, Shen X, Wang P. Constipation is associated with depression of any severity, but not with suicidal ideation: insights from a large cross-sectional study. Int J Colorectal Dis 2023;38(1):231. [doi: 10.1007/s00384-023-04520-8]
50. Liang J, Zhao Y, Xi Y, et al. Association between depression, anxiety symptoms and gut microbiota in Chinese elderly with functional constipation. Nutrients 2022 Nov 25;14(23):5013. [doi: 10.3390/nu14235013] [Medline: 36501044]
51. Costa LM, Mendes MM, Oliveira AC, et al. Dietary inflammatory index and its relationship with gut microbiota in individuals with intestinal constipation: a cross-sectional study. Eur J Nutr 2022 Feb;61(1):341-355. [doi: 10.1007/s00394-021-02649-2] [Medline: 34351455]
52. Turner-Stokes L, Frank AO. Urinary incontinence among patients with arthritis—a neglected disability. J R Soc Med 1992 Jul;85(7):389-393. [Medline: 1629846]
53. Fossaluzza V, Di Benedetto P, Zampa A, De Vita S. Misoprostol-induced urinary incontinence. J Intern Med 1991 Nov;230(5):463-464. [doi: 10.1111/j.1365-2796.1991.tb00474.x] [Medline: 1940783]
54. Mat S, Jaafar MH, Razack AHA, et al. Gender differences in the associations between knee pain and urinary incontinence in older adults: cross-sectional analysis from the Malaysian Elders Longitudinal Research Study (MELoR). Neurourol Urodyn 2023 Mar;42(3):641-649. [doi: 10.1002/nau.25136] [Medline: 36728321]
55. de Melo G, Fernandes VDO, Costa HS, et al. Urinary incontinence in middle-aged and older women with hip and knee osteoarthritis: an outpatient study of frequency and risk factors. Neurourol Urodyn 2024 Mar;43(3):672-679. [doi: 10.1002/nau.25395] [Medline: 38247352]
56. Judkins CP, Wang Y, Jelinic M, et al. Association of constipation with increased risk of hypertension and cardiovascular events in elderly Australian patients. Sci Rep 2023;13(1):10943. [doi: 10.1038/s41598-023-38068-y]
57. Davis NJ, Wyman JF, Gubitosa S, Pretty L. Urinary incontinence in older adults. Am J Nurs 2020;120(1):57-62. [doi: 10.1097/01.NAJ.0000652124.58511.24]
58. Camilleri M. Gastrointestinal motility disorders in neurologic disease. J Clin Invest 2021 Feb 15;131(4):e143771. [doi: 10.1172/JCI143771] [Medline: 33586685]
59. Erdogan T, Bahat G, Kilic C, et al. The relationship between sarcopenia and urinary incontinence. Eur Geriatr Med 2019 Dec;10(6):923-929. [doi: 10.1007/s41999-019-00232-x]
60. Li L, Chen F, Li X, Gao Y, Wang N, Xu T. Association between low handgrip strength and incontinence among Chinese middle-aged and older people: a cross-sectional study. Aging Med (Milton) 2024 Jun;7(3):360-367. [doi: 10.1002/agm2.12309] [Medline: 38975299]
61. Doumouchtsis SK, Loganathan J, Pergialiotis V. The role of obesity on urinary incontinence and anal incontinence in women: a review. BJOG 2022 Jan;129(1):162-170. [doi: 10.1111/1471-0528.16848] [Medline: 34312978]
62. Subak LL, Richter HE, Hunskaar S. Obesity and urinary incontinence: epidemiology and clinical research update. J Urol 2009 Dec;182(6 Suppl):S2-S7. [doi: 10.1016/j.juro.2009.08.071] [Medline: 19846133]
63. Zhu F, Chen M, Xiao Y, Huang X, Chen L, Hong L. Synergistic interaction between hyperlipidemia and obesity as a risk factor for stress urinary incontinence in Americans. Sci Rep 2024 Mar 27;14(1):7312. [doi: 10.1038/s41598-024-56744-5] [Medline: 38538650]

64. Suskind AM, Cawthon PM, Nakagawa S, et al. Urinary incontinence in older women: the role of body composition and muscle strength: from the health, aging, and body composition study. J Am Geriatr Soc 2017 Jan;65(1):42-50. [doi: 10.1111/jgs.14545] [Medline: 27918084]
65. Bauer SR, Grimes B, Suskind AM, Cawthon PM, Cummings S, Huang AJ. Urinary incontinence and nocturia in older men: associations with body mass, composition and strength in the health ABC study. Journal of Urology 2019 Nov;202(5):1015-1021. [doi: 10.1097/JU.0000000000000378]

Abbreviations

CHARLS: China Health and Retirement Longitudinal Study
OR: odds ratio
UDD: urinary and defecatory dysfunction

Edited by A Garcia; submitted 24.12.24; peer-reviewed by D Shaker, KH Goh; revised version received 30.03.25; accepted 17.04.25; published 26.05.25.

Please cite as:
Zhu H, Li X, Jing L, Kou J, Gong L, Wang R, Liu G, Zhang C, Zhao J, Zhang J
Factors Influencing the Severity of Urinary and Defecatory Dysfunction Among the Middle-Aged and Older Adult Chinese Population: Longitudinal Study of a 5-Wave Survey Cohort
JMIR Aging 2025;8:e70541
URL: https://aging.jmir.org/2025/1/e70541
doi:10.2196/70541

Original Paper

Predictive Validity of Hospital-Associated Complications of Older People Identified Using Diagnosis Procedure Combination Data From an Acute Care Hospital in Japan: Observational Study

Seigo Mitsutake[1,2], PhD, PT; Tatsuro Ishizaki[1], MD, PhD, MPH; Shohei Yano[1,3], PT, PhD; Takumi Hirata[1], MD, PhD, MPH; Kae Ito[1], MD, PhD; Ko Furuta[4], MD; Yoshitomo Shimazaki[5], RPh, PhD; Hideki Ito[6], MD, PhD; Alison Mudge[7,8], MD, PhD; Kenji Toba[6], MD, PhD

[1]Human Care Research Team, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[2]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[3]The Salvation Army Booth Memorial Hospital, Tokyo, Japan

[4]Department of Psychiatry, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[5]Department of Pharmacy, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[6]Tokyo Metropolitan Institute for Geriatrics and Gerontology, Tokyo, Japan

[7]Department of Internal Medicine and Aged Care, Royal Brisbane and Women's Hospital, Brisbane, Australia

[8]Centre of Health Services Research, The University of Queensland, Brisbane, Australia

Corresponding Author:
Seigo Mitsutake, PhD, PT
Human Care Research Team
Tokyo Metropolitan Institute for Geriatrics and Gerontology
35-2 Sakae-cho
Itabashi-ku
Tokyo, 173 0015
Japan
Phone: 81 3 3964 3241 ext 4229
Email: mitsu@tmig.or.jp

Abstract

Background: A composite outcome of hospital-associated complications of older people (HAC-OP; comprising functional decline, delirium, incontinence, falls, and pressure injuries) has been proposed as an outcome measure reflecting quality of acute hospital care. Estimating HAC-OP from routinely collected administrative data could facilitate the rapid and standardized evaluation of interventions in the clinical setting, thereby supporting the development, improvement, and wider implementation of effective interventions.

Objective: This study aimed to create a Diagnosis Procedure Combination (DPC) data version of the HAC-OP measure (HAC-OP-DPC) and demonstrate its predictive validity by assessing its associations with hospital length of stay (LOS) and discharge destination.

Methods: This retrospective cohort study acquired DPC data (routinely collected administrative data) from a general acute care hospital in Tokyo, Japan. We included data from index hospitalizations for patients aged ≥65 years hospitalized for ≥3 days and discharged between July 2016 and March 2021. HAC-OP-DPC were identified using diagnostic codes for functional decline, incontinence, delirium, pressure injury, and falls occurring during the index hospitalization. Generalized linear regression models were used to examine the associations between HAC-OP-DPC and LOS, and logistic regression models were used to examine the associations between HAC-OP-DPC and discharge to other hospitals and long-term care facilities (LTCFs).

Results: Among 15,278 patients, 3610 (23.6%) patients had coding evidence of one or more HAC-OP-DPC (1: 18.8% and ≥2: 4.8%). Using "no HAC-OP-DPC" as the reference category, the analysis showed a significant and graded association with longer LOS (adjusted risk ratio for patients with one HAC-OP-DPC 1.29, 95% CI 1.25-1.33; adjusted risk ratio for ≥2 HAC-OP-DPC 1.97, 95% CI 1.87-2.08), discharge to another hospital (adjusted odds ratio [AOR] for one HAC-OP-DPC 2.36, 95% CI 2.10-2.65; AOR for ≥2 HAC-OP-DPC 6.96, 95% CI 5.81-8.35), and discharge to LTCFs (AOR for one HAC-OP-DPC 1.35, 95% CI

1.09-1.67; AOR for ≥2 HAC-OP-DPC 1.68, 95% CI 1.18-2.39). Each individual HAC-OP was also significantly associated with longer LOS and discharge to another hospital, but only delirium was associated with discharge to LTCF.

Conclusions: This study demonstrated the predictive validity of the HAC-OP-DPC measure for longer LOS and discharge to other hospitals and LTCFs. To attain a more robust understanding of these relationships, additional studies are needed to verify our findings in other hospitals and regions. The clinical implementation of HAC-OP-DPC, which is identified using routinely collected administrative data, could support the evaluation of integrated interventions aimed at optimizing inpatient care for older adults.

(JMIR Aging 2025;8:e68267) doi:10.2196/68267

KEYWORDS

delirium; functional decline; Japan; older adult; routinely collected health data; elder; hospital complication; HAC-OP; incontinence; pressure injury; inpatient care; diagnosis procedure combination; predictive validity; hospital length of stay; administrative data; acute care; index hospitalization; diagnostic code; linear regression; logistic regression; long-term care; retrospective cohort; observational study; patient care; gerontology; hospital care; patient complication

Introduction

Hospitalization can impose a heavy physical and psychological burden on older adults, leading to in-hospital complications such as functional decline (the loss of independence in activities of daily living) and delirium [1]. A recent meta-analysis reported that the prevalence of functional decline following acute hospitalization was 30% in older adults aged ≥65 years [2]. Such functional decline is associated with various adverse outcomes, including prolonged hospital length of stay (LOS) and increased risks of hospital readmission and mortality within 30 days after discharge [3,4]. Similarly, delirium is estimated to affect around one quarter of older hospitalized adults [5-7], and is associated with longer stays, higher mortality, and higher health care costs [8,9].

Previous studies have developed and demonstrated the effectiveness of interventions to prevent such complications in hospitalized older adults [10,11]. While these interventions have generally focused on single complications as outcomes, they may also be effective in preventing other complications (eg, delirium prevention programs can also reduce falls) [1,12-16]. This is because in-hospital complications common in older people (often termed “geriatric syndromes”) tend to have overlapping risk factors that can be collectively targeted by complex and multifaceted interventions [1,12-16]. Mudge et al [1] proposed a multicomponent measure of “hospital-associated complications of older people” (HAC-OP) comprising common complications (functional decline, hospital-associated incontinence, hospital-associated delirium, pressure injury, or fall) among older adults admitted to acute care hospitals. In a prospective study using regular structured patient assessments as well as document review, the authors demonstrated a significant graded association between the HAC-OP measure and hospital LOS, facility discharge, and mortality within 6 months after admission. However, collecting this research measure was resource-intensive [1,17], which makes it difficult to replicate in routine practice. Routinely collected administrative data have been examined to assist clinicians, patients, and policy makers in making informed decisions [18,19]. Estimating HAC-OP from routinely collected administrative data could facilitate the rapid and standardized evaluation of interventions in the clinical setting, thereby supporting the development, improvement, and wider implementation of effective interventions.

In Japan, the majority of acute care hospitals use the Diagnosis Procedure Combination (DPC) case-mix patient classification system, which is linked to a lump-sum payment system for inpatients [20]. DPC-compliant hospitals must generate and submit DPC data to the government. These data include administrative claims and discharge abstracts containing patient-level information on diagnoses (recorded using *ICD-10* [*International Statistical Classification of Diseases, Tenth Revision*] codes), treatments, and prescribed drugs. Unlike many other countries, the required data also include activities of daily living (ADL) assessment scores at admission and discharge [20]. Prior studies have used DPC data to assess ADL scores in patients to examine the effects of rehabilitation services and to identify patients at high risk of early readmission in acute care settings [21,22]. Hospital outcomes, such as LOS and readmission within 28 days after discharge from hospital, are published annually for each DPC hospital [23]. However, in-hospital complications are not currently reported or benchmarked as they are in some other countries [24,25]. The systematic and accurate identification of HAC-OP from DPC data would support the evaluation of inpatient care and interventions that target these complications in Japan’s acute care hospitals.

Although the validity of chronic disease diagnoses in DPC data has been reported to be generally high [26], no study has tested the validity of in-hospital complications in DPC data, and complications are often underestimated in administrative datasets [27]. Understanding the current reporting of in-hospital complications and the association with important outcomes such as LOS and discharge destination could inform their use as an efficient and standardized evaluation of system-level interventions. Therefore, this study was conducted to develop a DPC data version of the HAC-OP measure (HAC-OP-DPC), describe the incidence of HAC-OP-DPC in a cohort of older acute care inpatients, and evaluate the predictive validity of this composite measure and its components by assessing associations with LOS and discharge destination.

Methods

Study Design and Patients

This retrospective cohort study was conducted using an anonymized DPC database obtained from a large general public acute care hospital (550 beds: 520 beds on general wards and 30 beds on psychiatric wards) in Tokyo, Japan. The DPC data comprised patient-level demographic characteristics, *ICD-10* codes, treatments, and prescribed drugs during all insurance-covered clinical encounters. The study used data from July 2016 to March 2021.

Patients who had been admitted to the study hospital from home or a long-term care facility (LTCF) and discharged during the study period were eligible for inclusion; the first hospitalization episode during the study period was designated the index hospitalization and included in the analysis. We excluded patients aged <65 years, patients discharged within 2 days of admission, patients who died during the index hospitalization, and patients with missing data in the study variables.

Ethical Considerations

The study protocol was approved by the Ethics Committee of the Tokyo Metropolitan Geriatric Hospital and Institute of Gerontology (approval number R18-20). All procedures followed the ethical guidelines of the Medical and Biological Research Involving Human Subjects established by the Japanese government. Opt-out consent was used because all data were anonymized before being received by the authors.

Measures

HAC-OP-DPC

Based on the original HAC-OP measure [1], we identified the following 5 conditions as HAC-OP-DPC: hospital-associated functional decline, incontinence, delirium, pressure injury, and fall. Each condition was defined using DPC data.

First, hospital-associated functional decline was defined as a decrease in ADL scores for the Barthel Index (BI) components of bathing, dressing, toileting, transfers, mobility, and feeding from hospital admission to discharge. In the BI, each component is given a score of 0, 5, 10, or 15 points (maximum scores vary among the components), with higher scores indicating greater independence in that activity [28]. BI was assessed at admission and hospital discharge by bedside nurses. Second, hospital-associated incontinence was defined as a decrease in scores for the BI components of bladder function and bowel function from hospital admission to discharge. Third, hospital-associated delirium was identified based on a recorded diagnosis of delirium (Multimedia Appendix 1) as a postadmission complication and recorded prescriptions of drugs used to manage agitation in delirium (injections of haloperidol or other antipsychotic drugs identified using prescription codes that remained constant throughout the study period). Fourth, hospital-associated pressure injury was identified based on a recorded diagnosis of pressure injury as a postadmission complication and discharge abstract records indicating pressure injury treatment during hospitalization without any similar treatment at admission. Fifth, a hospital-associated fall was identified based on a recorded diagnosis of fall as a postadmission complication. The overall multicomponent HAC-OP-DPC measure was categorized into none, 1, and 2 or more complications based on the count of conditions occurring in each patient.

Outcome Measures

The study outcome measures were hospital LOS during the index hospitalization, discharge to other hospitals, and discharge to LTCFs. LOS was calculated as the number of days between the dates of admission and discharge. Discharges to other hospitals, such as rehabilitation hospitals and LTCFs (including special nursing homes, private paid care facilities, and social welfare institutions), were identified using the relevant DPC codes indicating discharge destination.

Covariates

Using the subject hospital's DPC data, we extracted demographic variables of patient sex, age group (65-74, 75-84, and ≥85 years), and annual household income (<¥3.7 million, ≥¥3.7 million, and unknown; ¥1=US $0.0092 in 2016) at the index hospitalization [29]. Income was estimated from the available data about insurance copayments. Insurance copayment rates are the designated rates that patients pay at the point of care in Japan. For patients who have an annual household income below ¥3.7 million (approximately US $34,040; ¥1=US $0.0092), the copayment rates are 10% and 20% for patients aged ≥75 years and 70-74 years, respectively [29]. For patients aged ≥70 years who have an annual household income of ¥3.7 million or higher (¥1=US $0.0092), the copayment rate is 30%. The DPC data did not indicate the copayment rates for patients who received public medical assistance and patients aged 65-69 years. Therefore, income for these cases was categorized as "unknown."

We calculated variables for disease category, comorbidity, and frailty using *ICD-10* codes. Using previously described methods [30,31], we grouped patients into 12 disease categories based on their recorded primary diagnosis for admission. Next, we determined each patient's score in the Charlson Comorbidity Index (CCI), which is a weighted index of specific comorbidities that were identified using *ICD-10* codes [32]. CCI scores were divided into 3 categories (0, 1-2, and ≥3). Similarly, we calculated each patient's Hospital Frailty Risk Score (HFRS), which was developed to identify older adults experiencing frailty with a higher risk of adverse outcomes [33]. The total HFRS ranges from 0 to 99 and was divided into 2 categories (<5 and ≥5). To determine baseline functional dependence and incontinence levels in patients at admission, we analyzed the following 2 variables: dependence in ≥1 ADL items (BI components of bathing, dressing, toileting, transfers, mobility, and feeding) at admission and urinary and fecal incontinence (BI components of bladder function and bowel function) at admission. We dichotomized each of these 2 variables into independent (ie, scoring the maximum score on all components) or dependent (all other patients). We also determined each patient's location before admission (home or LTCF) and the surgical treatment received during the index hospitalization.

Statistical Analysis

The chi-square test was used to compare the differences in patient characteristics among the 3 HAC-OP-DPC categories. We generated cross-tabulations to examine the co-occurrences of each complication. Pearson correlation coefficients were calculated to measure the associations between each complication.

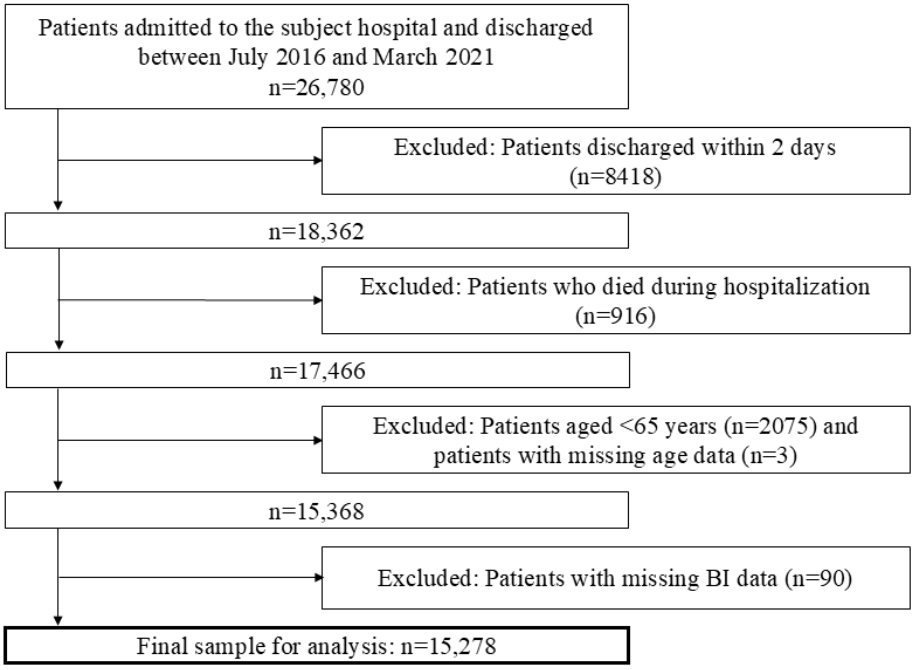
The associations between HAC-OP-DPC and LOS were examined using multivariable generalized linear regression models for gamma-distributed data with a log-link function that adjusted for all covariates. Effect sizes for 1 HAC-OP-DPC and ≥2 HAC-OP-DPC were quantified using adjusted risk ratios (ARRs) and their 95% CIs, which indicated the likelihood of having a longer LOS. Next, the associations of HAC-OP-DPC with discharge to other hospitals and discharge to LTCFs were examined using multivariable logistic regression models that adjusted for all covariates. Effect sizes for 1 HAC-OP-DPC and ≥2 HAC-OP-DPC were quantified using adjusted odds ratios (AORs) and their 95% CIs, which indicated the odds of being discharged to another hospital or LTCF. In addition to analyzing the associations between the number of HAC-OP-DPC and the 3 outcomes, we constructed models to examine the associations between individual HAC-OP-DPC components and the outcomes. For the analysis of hospital-associated functional decline, we excluded patients who were already dependent in all the BI components at admission because they could not experience any further functional decline. Similarly, for the analysis of hospital-associated incontinence, we excluded patients who were already dependent in both bladder function and bowel function at admission because they could not experience new-onset incontinence during hospitalization. We also conducted several sensitivity analyses. First, we re-ran the primary analysis for each of the 3 outcomes, excluding patients who could not experience functional decline (fully dependent on admission) or hospital-associated incontinence (incontinent on admission). Second, we recalculated the HAC-OP-DPC composite outcome excluding cases of delirium that were identified only by drug prescribing as we recognized that antipsychotic drugs might be prescribed for other indications (eg, behavioral and psychological symptoms of dementia, psychosis). All analyses were conducted using SPSS (version 28.0; IBM Corp), and the cross-tabulations were created using R (version 4.3.2; R project for Statistical Computing). *P* values (2-tailed) below .05 were considered statistically significant.

Results

Study Participant Selection

We first identified 26,780 candidate patients who were admitted to the subject hospital and discharged to home, another hospital, or an LTCF during the study period (Figure 1). After applying the exclusion criteria, the final sample for analysis consisted of 15,278 patients aged ≥65 years. Multimedia Appendix 2 shows the characteristics among 833 patients ≥65 years who died during hospitalization and without missing data at admission. These excluded participants were aged ≥85 years, had ≥2 CCI, were more likely to have ADL dependency and incontinence, and were more likely to have cancer, pneumonia, or heart failure.

Figure 1. Flowchart of patient selection. BI: Barthel Index.



Characteristics of Study Patients

Their mean age was 81.2 (SD 7.9) years, and women accounted for 55.4% of all patients (Table 1). There were 2877 (18.8%) patients who experienced 1 HAC-OP-DPC and 733 (4.8%) patients who experienced ≥2 HAC-OP-DPC during the index hospitalization.

Table 1. Patient characteristics according to the number of HAC-OP-DPC[a].

Characteristics	Total (N=15,278)		HAC-OP-DPC						*P* value
			None (n=11,668)		1% (n=2877)		≥2% (n=733)		
	N	%	n	%	n	%	n	%	
Sex									.16
Men	6815	44.6	5183	44.4	1280	44.5	352	48.0	
Women	8463	55.4	6485	55.6	1597	55.5	381	52.0	
Age (years)									<.001
65-74	3353	21.9	2771	23.7	508	17.7	74	10.1	
75-84	6396	41.9	5009	42.9	1117	38.8	270	36.8	
≥85	5529	36.2	3888	33.3	1252	43.5	389	53.1	
Annual household income (¥; ¥ 1=US $0.0092)									<.001
<3.7 million	11,600	75.9	8759	75.1	2230	77.5	611	83.4	
≥3.7 million	1570	10.3	1195	10.2	308	10.7	67	9.1	
Unknown	2108	13.8	1714	14.7	339	11.8	55	7.5	
Primary diagnosis for admission									<.001
Musculoskeletal diseases	1127	7.4	890	7.6	203	7.1	34	4.6	
Coronary heart disease	338	2.2	288	2.5	47	1.6	3	0.4	
Congestive heart failure	689	4.5	505	4.3	133	4.6	51	7.0	
Cerebrovascular disease	1039	6.8	874	7.5	133	4.6	32	4.4	
Pneumonia or acute bronchitis	1323	8.7	974	8.3	274	9.5	75	10.2	
Fracture	576	3.8	407	3.5	137	4.8	32	4.4	
Metabolic diseases	972	6.4	735	6.3	186	6.5	51	7.0	
Renal diseases	879	5.8	646	5.5	190	6.6	43	5.9	
Neurological diseases	631	4.1	510	4.4	100	3.5	21	2.9	
Gastrointestinal diseases	1694	11.1	1272	10.9	352	12.2	70	9.5	
Cancer	1867	12.2	1344	11.5	400	13.9	123	16.8	
Other	4143	27.1	3223	27.6	722	25.1	198	27.0	
CCI[b]									<.001
0	9041	59.2	7116	61.0	1563	54.3	362	49.4	
1-2	4852	31.8	3578	30.7	1000	34.8	274	37.4	
≥3	1385	9.1	974	8.3	314	10.9	97	13.2	
HFRS[c]									<.001
<5	13,278	86.9	10,324	88.5	2371	82.4	583	79.5	
≥5	2000	13.1	1344	11.5	506	17.6	150	20.5	
Dependence in ≥1 ADL[d] items at admission	12,658	82.9	9764	83.7	2239	77.8	655	89.4	<.001
Urinary and fecal incontinence at admission	5158	33.8	3932	33.7	1038	36.1	188	25.6	<.001
Location before admission									<.001
Home	14,186	92.9	10,875	93.2	2642	91.8	669	91.3	
LTCF[e]	1092	7.1	793	6.8	235	8.2	64	8.7	
Surgical treatment	3337	21.8	2444	20.9	732	25.4	161	22.0	<.001

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.

[b]CCI: Charlson Comorbidity Index.
[c]HFRS: Hospital Frailty Risk Score.
[d]ADL: activities of daily living.
[e]LTCF: long-term care facility.

HAC-OP-DPC Among This Study's Patients

The most common complication was functional decline (n=2103, 13.8%), followed by delirium (n=1345, 8.8%: recorded diagnosis of delirium (n=59); recorded prescriptions of drugs for delirium, n=1286), new incontinence (n=860, 5.6%), pressure injury (n=104, 0.7%), and fall (n=59, 0.4%). Figure 2 demonstrates the patterns of co-occurrence amongst HAC-OP-DPC. For example, among patients with functional decline, 20.6% also experienced incontinence during hospitalization, while 50.5% of patients with incontinence, 33.9% of patients with falls, and 22.5% of patients with delirium also experienced functional decline during hospitalization. While there were statistically significant correlations between most complications, the strength of these correlations was weak (maximum correlation coefficient of 0.26; Table 2). As shown in Table 1, older age, annual household income, multimorbidity, frailty, baseline functional impairment and incontinence, and living in LTCF were all significantly associated with HAC-OP-DPC. Patients with ≥2 HAC-OP-DPC had a significantly greater proportion of patients with congestive heart failure, pneumonia or acute bronchitis, metabolic diseases, and cancer as the principal diagnosis.

Figure 2. Co-occurrence of each HAC-OP-DPC among older patients aged ≥65 years (N=15,278). Values show the percentages of patients in each row who also had the column condition during hospitalization. HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.

Table 2. Within-patient correlations between each HAC-OP-DPC[a] among older patients aged ≥65 years (N=15,278). Values show Pearson correlation coefficients (r) between each HAC-OP-DPC.

	Hospital-associated functional decline	Hospital-associated incontinence	Hospital-associated delirium	Hospital-associated pressure injury	Hospital-associated fall
Hospital-associated functional decline	1.000	0.260[b]	0.079[b]	–0.008	0.036[b]
Hospital-associated incontinence	—	1.000	0.099[b]	0.021[b]	0.040[b]
Hospital-associated delirium	—	—	1.000	0.016[c]	0.025[b]
Hospital-associated pressure injury	—	—	—	1.000	0.008
Hospital-associated fall	—	—	—	—	1.000

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.
[b]$P<.01$.
[c]$P<.05$.

Association of HAC-OP-DPC With Outcome Measures

Table 3 demonstrates the significant and graded association observed between HAC-OP-DPC and outcomes. Adjusting for all covariates, those with 1 and ≥2 HAC-OP-DPC were significantly more likely to have a longer LOS during the index hospitalization (ARR 1.29, 95% CI 1.25-1.33, and ARR 1.97, 95% CI 1.87-2.08, respectively) compared to those with no HAC-OP-DPC. They were significantly more likely to be discharged to other hospitals (AOR 2.36, 95% CI 2.10-2.65, and AOR 6.96, 95% CI 5.81-8.35, respectively), and be discharged to LTCFs (AOR 1.35, 95% CI 1.09-1.67, and AOR 1.68, 95% CI 1.18-2.39, respectively). The analyses of individual HAC-OP-DPC showed that patients who experienced hospital-associated functional decline, incontinence, delirium, pressure injury, and falls were significantly more likely to have a longer LOS and be discharged to other hospitals than patients without these complications. However, only delirium was significantly associated in this individual complication analysis with discharge to LCTFs. A sensitivity analysis (n=11,075) that excluded 4203 patients who could neither experience hospital-associated functional decline nor incontinence found that HAC-OP-DPC still had a significant and graded association with longer LOS and discharge destination (Multimedia Appendix 3), similar to the results of the main analysis. Moreover, the associations of the number of HAC-OP-DPC using delirium identified from a recorded diagnosis alone with these outcomes were similar to the results of the main analysis (Multimedia Appendix 4).

Table 3. Association of HAC-OP-DPCa with outcome measures in 15,278 participants.

	Participants, n	LOS[b]				Discharge to other hospitals			Discharge to LTCFs[c]		
		Median	IQR	RR[d] (95% CI)[e]	ARR[f] (95% CI)[g]	%	OR[h] (95% CI)[i]	AOR[j] (95% CI)[k]	%	OR (95% CI)[i]	AOR (95% CI)[k]
HAC-OP-DPC											
No (Ref[l])	11,482	14	(7-25)	1.00	1.00	14.4	1.00	1.00	7.3	1.00	1.00
1	2,838	18	(9.5-32)	1.28 (1.25-1.32)	1.29 (1.25-1.33)	24.6	1.94 (1.76-2.14)	2.36 (2.10-2.65)	9.7	1.42 (1.24-1.64)	1.35 (1.09-1.67)
≥2	724	29	(14-44)	1.88 (1.78-1.99)	1.97 (1.87-2.08)	42.6	4.40 (3.76-5.14)	6.96 (5.81-8.35)	11.1	1.76 (1.38-2.23)	1.68 (1.18-2.39)
Hospital-associated functional decline[m]											
No (Ref)	10,124	13	(7-24)	1.00	1.00	10.4	1.00	1.00	4.3	1.00	1.00
Yes	2,103	16	(8-31)	1.35 (1.30-1.40)	1.35 (1.30-1.40)	23.6	2.67 (2.37-3.00)	3.56 (3.11-4.09)	7.3	1.77 (1.46-2.14)	1.09 (0.81-1.47)
Hospital-associated incontinence[n]											
No (Ref)	10,615	12	(7-22)	1.00	1.00	8.6	1.00	1.00	3.4	1.00	1.00
Yes	860	27	(14-42)	1.93 (1.83-2.04)	1.80 (1.71-1.9)	43.3	8.12 (6.98-9.44)	7.33 (6.16-8.73)	13.3	4.39 (3.51-5.49)	1.24 (0.84-1.83)
Hospital-associated delirium											
No (Ref)	13,933	14	(8-26)	1.00	1.00	16.1	1.00	1.00	7.6	1.00	1.00
Yes	1,345	25	(14-40.5)	1.61 (1.54-1.68)	1.53 (1.47-1.59)	34.1	2.70 (2.39-3.05)	2.23 (1.94-2.56)	12.3	1.71 (1.44-2.04)	1.61 (1.24-2.10)
Hospital-associated pressure injury											
No (Ref)	15,174	15	(8-27)	1.00	1.00	17.5	1.00	1.00	7.9	1.00	1.00
Yes	104	28	(16.25-42.75)	1.68 (1.45-1.95)	1.43 (1.24-1.64)	48.1	4.37 (2.97-6.43)	2.78 (1.83-4.22)	17.3	2.44 (1.46-4.07)	0.69 (0.31-1.56)
Hospital-associated fall											
No (Ref)	15,219	15	(8-27)	1.00	1.00	17.6	1.00	1.00	8.0	1.00	1.00
Yes	59	38	(22-50)	2.05 (1.68-2.49)	2.02 (1.68-2.43)	49.2	4.53 (2.72-7.56)	4.65 (2.63-8.23)	5.1	0.62 (0.19-1.98)	0.68 (0.15-3.19)

[a]HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version.

[b]LOS: length of stay.

[c]LTCF: long-term care facility.

[d]RR: risk ratio.

[e]Generalized linear regression analysis.

[f]ARR: adjusted risk ratio.

[g]Generalized linear regression analysis that adjusted for all covariates (sex, age group, annual household income, primary diagnosis for admission, Charlson Comorbidity Index score, Hospital Frailty Risk Score, dependence in ≥1 activities of daily living items at admission, urinary and fecal incontinence at admission, location before admission, and surgical treatment).

[h]OR: odds ratio.

[i]Logistic regression analysis.

[j]AOR: adjusted odds ratio.

[k]Logistic regression analysis that adjusted for all covariates (sex, age group, annual household income, primary diagnosis for admission, Charlson Comorbidity Index score, Hospital Frailty Risk Score, dependence in ≥1 activities of daily living items at admission, urinary and fecal incontinence at

admission, location before admission, and surgical treatment).

[l]Ref: reference.

[m]Models for functional decline excluded participants with pre-exiting full dependence (model n=12,227).

[n]Models for hospital-associated incontinence excluded those with pre-existing incontinence (model n=11,475).

Discussion

Principal Findings

This retrospective cohort study is the first to develop and apply a tool to assess HAC-OP from routinely collected administrative data and evaluate its predictive validity for hospital outcomes. Our analysis showed that almost one quarter of older inpatients with multiday stays of more than 2 days had a coded HAC-OP, and that having one or more HAC-OP-DPC was associated with longer LOS and discharge to other hospitals and LTCFs. The clinical implementation of the HAC-OP-DPC measure could support comparative analyses of clinical and policy interventions aimed at preventing these complications, thereby contributing to the optimization of acute care for older adults.

Almost 1 in 4 older patients had coding documentation of any HAC-OP-DPC in the present study, which was approximately half that of the incidence of any HAC-OP described in the Australian study [1]. Nonetheless, the incidences of each HAC-OP-DPC were ranked in similar order, with functional decline and delirium being most common. A large contribution to the disparity in incidence is likely to be reliance on DPC data to identify individual complications as HAC-OP-DPC, while the HAC-OP study used repeated patient and clinical record assessments by trained research assistants [1,17]. Although we made efforts to minimize the underestimation of delirium, falls, and pressure injuries by including additional codes and data sources in their identification criteria, it is very likely that the coding data underestimated the incidences of all complications. For example, there were no coded falls without fracture, and only 59 cases of direct recording of a delirium diagnosis. There are recognized gaps in clinician recognition and documentation of hospital complications as well as translation into coding [27]. Understanding and improving the accuracy and usability of HAC-OP-DPC may require correlation with clinical data and comparisons between sites. There are other reasons that our incidence estimates may have been lower than expected. The denominator in our main analysis included 3051 patients who could not experience further functional decline and 3803 patients who could not experience new-onset incontinence, which may have led to an underestimation of these HAC-OP-DPC. Our analyses also excluded patients who died during their inpatient stay, who may have had a higher rate of HAC-OP.

Nonetheless, our study found that the HAC-OP-DPC measure was associated with longer LOS and discharge destination, thereby demonstrating its predictive validity for outcomes in an acute care setting. Furthermore, our analysis found nonoverlapping risk estimates for LOS and discharge to other hospitals between patients with 1 and ≥2 HAC-OP-DPC, showing a significant exposure-outcome effect of a graded nature. These findings were consistent with the results of the original HAC-OP study [1] and suggest that prevention of further HAC-OP is important in those who have already acquired 1 complication. Our observations that HAC-OP-DPC was significantly associated with older age and baseline function are consistent with previous reports [1]. Importantly, our study is the first to demonstrate the strong associations of HAC-OP with higher CCI scores and HFRS, consistent with existing knowledge that comorbidities and frailty are risk factors for the individual complications included in the composite measure [13,16,34-37].

Our analysis also confirmed that the individual HAC-OP-DPC were associated with longer LOS and discharge to other hospitals, which was congruent with previous studies [6,14,35-39]. All HAC-OP-DPC have been individually recognized as important outcomes in older patients [1,6,14,17,36-42]. Although our study showed that there were patterns of co-occurrence and significant correlations among the individual HAC-OP-DPC, these correlations were weak. This suggests that these complications represent relatively distinct conditions and that they can be treated individually [1]. Moreover, a systematic review of composite outcomes in clinical trials proposed that studies should list results for all components of a composite outcome to avoid confusion and bias [43], and an outcome study using the original multicomponent HAC-OP measure demonstrated significant and clinically important improvements in individual outcomes but not in the composite measure [17]. We recommend that future studies should report the effects of interventions not only on the HAC-OP-DPC measure as a composite outcome but also on its individual complications.

Strengths and Limitations

Strengths of this study include a large, representative dataset with high levels of item completeness (including functional variables), adjustment for important covariates, and use of sensitivity analyses to explore data assumptions. We also recognize several limitations. First, we included additional information to reduce these anticipated underestimates that may have reduced precision; for example, by including drug prescribing of antipsychotics in the delirium diagnosis, we may have included some patients with other indications such as behavioral and psychological symptoms of dementia. However, our sensitivity analysis using a more stringent definition of delirium suggests that this had minimal impact on our overall findings. Second, our study was conducted in a single acute care hospital in Japan, and its findings may not be generalizable to other hospitals, regions, or countries. Nevertheless, the majority of acute care hospitals in Japan have adopted the DPC system [20], and future studies could compare the incidence of HAC-OP-DPC in each acute hospital throughout Japan. Also, although a previous study identified the validity of chronic disease diagnoses in DPC data as being generally high [26], we recognize that the performance of comorbidity scoring methods based on administrative data may vary between health systems [44]. Third, we could not infer a causal relationship between HAC-OP-DPC and LOS because we were unable to identify the date when each complication occurred, and we cannot

exclude reverse causality between HAC-OP-DPC and LOS [1]. Fourth, our study did not examine the association between HAC-OP-DPC and mortality, which had been analyzed in the original HAC-OP study [1], because there are no BI scores assigned on discharge for in-hospital deaths to calculate functional decline and incontinence of HAC-OP-DPC, and DPC data lack information on death after discharge. Future studies could link DPC data with mortality data to examine the associations between HAC-OP-DPC and mortality. Finally, our database lacked information on potentially important confounders, such as residential status (eg, living alone or with others), the presence of caregivers, and disease severity that could be included in future studies.

Conclusions

This study showed that almost one quarter of older acute care inpatients in a Japanese hospital have coding indicating a HAC-OP and demonstrated the predictive validity of the HAC-OP-DPC measure for longer LOS and discharge to other hospitals and LTCFs. To attain a more robust understanding of these relationships, additional studies are needed to verify our findings in other hospitals and regions. The clinical implementation of HAC-OP-DPC, which are identified using routinely collected administrative data, could support the efficient evaluation of integrated interventions aimed at optimizing inpatient care for older adults.

Acknowledgments

This study was supported by KAKENHI Grants-in-Aid for Scientific Research (grant numbers 20H03924 and 24K13306) from the Japan Society for the Promotion of Science.

Conflicts of Interest

None declared.

Multimedia Appendix 1
ICD-10 (International Statistical Classification of Diseases, Tenth Revision) codes for identifying hospital-associated complications of older people-Diagnosis Procedure Combination data version.
[DOCX File , 34 KB - aging_v8i1e68267_app1.docx]

Multimedia Appendix 2
Characteristics among patients who died during hospitalization (n=833).
[DOCX File , 36 KB - aging_v8i1e68267_app2.docx]

Multimedia Appendix 3
Association of HAC-OP-DPC (hospital-associated complications of older people-Diagnosis Procedure Combination data version) with the outcome measures after excluding patients who could neither experience hospital-associated functional decline nor incontinence (n=11,075).
[DOCX File , 36 KB - aging_v8i1e68267_app3.docx]

Multimedia Appendix 4
Association of HAC-OP-DPC (hospital-associated complications of older people-Diagnosis Procedure Combination data version) with the outcome measures after excluding patients who could neither experience hospital-associated functional decline nor incontinence (N=15,278).
[DOCX File , 36 KB - aging_v8i1e68267_app4.docx]

References

1. Mudge AM, McRae P, Hubbard RE, Peel NM, Lim WK, Barnett AG, et al. Hospital-associated complications of older people: a proposed multicomponent outcome for acute care. J Am Geriatr Soc 2019 Feb;67(2):352-356 [FREE Full text] [doi: 10.1111/jgs.15662] [Medline: 30423197]
2. Loyd C, Markland AD, Zhang Y, Fowler M, Harper S, Wright NC, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020;21(4):455-461.e5 [FREE Full text] [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
3. Boyd CM, Ricks M, Fried LP, Guralnik JM, Xue Q, Xia J, et al. Functional decline and recovery of activities of daily living in hospitalized, disabled older women: the women's health and aging study I. J Am Geriatr Soc 2009;57(10):1757-1766 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02455.x] [Medline: 19694869]
4. Berian JR, Mohanty S, Ko CY, Rosenthal RA, Robinson TN. Association of loss of independence with readmission and death after discharge in older patients after surgical procedures. JAMA Surg 2016;151(9):e161689. [doi: 10.1001/jamasurg.2016.1689] [Medline: 27409710]

5. Gibb K, Seeley A, Quinn T, Siddiqi N, Shenkin S, Rockwood K, et al. The consistent burden in published estimates of delirium occurrence in medical inpatients over four decades: a systematic review and meta-analysis study. Age Ageing 2020;49(3):352-360 [FREE Full text] [doi: 10.1093/ageing/afaa040] [Medline: 32239173]
6. Geriatric Medicine Research Collaborative. Delirium is prevalent in older hospital inpatients and associated with adverse outcomes: results of a prospective multi-centre study on world delirium awareness day. BMC Med 2019;17(1):229 [FREE Full text] [doi: 10.1186/s12916-019-1458-7] [Medline: 31837711]
7. Siddiqi N, House A, Holmes J. Occurrence and outcome of delirium in medical in-patients: a systematic literature review. Age Ageing 2006;35(4):350-364. [doi: 10.1093/ageing/afl005] [Medline: 16648149]
8. Milbrandt EB, Deppen S, Harrison PL, Shintani AK, Speroff T, Stiles RA, et al. Costs associated with delirium in mechanically ventilated patients. Crit Care Med 2004;32(4):955-962. [doi: 10.1097/01.ccm.0000119429.16055.92] [Medline: 15071384]
9. Leslie DL, Marcantonio ER, Zhang Y, Leo-Summers L, Inouye SK. One-year health care costs associated with delirium in the elderly population. Arch Intern Med 2008;168(1):27-32 [FREE Full text] [doi: 10.1001/archinternmed.2007.4] [Medline: 18195192]
10. Cohen Y, Zisberg A, Chayat Y, Gur-Yaish N, Gil E, Levin C, et al. Walking for better outcomes and recovery: the effect of WALK-FOR in preventing hospital-associated functional decline among older adults. J Gerontol A Biol Sci Med Sci 2019;74(10):1664-1670. [doi: 10.1093/gerona/glz025] [Medline: 30726886]
11. Clinical guidelines, No. 103. Delirium: prevention, diagnosis and management in hospital and long-term care. National Institute for Health and Care Excellence (NICE). London: (NICE) URL: https://www.nice.org.uk/guidance/cg103 [accessed 2024-03-18]
12. Hshieh TT, Yue J, Oh E, Puelle M, Dowal S, Travison T, et al. Effectiveness of multicomponent nonpharmacological delirium interventions: a meta-analysis. JAMA Intern Med 2015;175(4):512-520 [FREE Full text] [doi: 10.1001/jamainternmed.2014.7779] [Medline: 25643002]
13. Geyskens L, Jeuris A, Deschodt M, Van Grootven B, Gielen E, Flamaing J. Patient-related risk factors for in-hospital functional decline in older adults: a systematic review and meta-analysis. Age Ageing 2022;51(2):afac007. [doi: 10.1093/ageing/afac007] [Medline: 35165688]
14. Ahmed S, Leurent B, Sampson EL. Risk factors for incident delirium among older people in acute hospital medical units: a systematic review and meta-analysis. Age Ageing 2014;43(3):326-333 [FREE Full text] [doi: 10.1093/ageing/afu022] [Medline: 24610863]
15. Deandrea S, Bravi F, Turati F, Lucenteforte E, La Vecchia C, Negri E. Risk factors for falls in older people in nursing homes and hospitals. A systematic review and meta-analysis. Arch Gerontol Geriatr 2013;56(3):407-415. [doi: 10.1016/j.archger.2012.12.006] [Medline: 23294998]
16. Palmer MH, Baumgarten M, Langenberg P, Carson JL. Risk factors for hospital-acquired incontinence in elderly female hip fracture patients. J Gerontol A Biol Sci Med Sci 2002;57(10):M672-M677. [doi: 10.1093/gerona/57.10.m672] [Medline: 12242323]
17. Mudge AM, McRae P, Banks M, Blackberry I, Barrimore S, Endacott J, et al. Effect of a ward-based program on hospital-associated complications and length of stay for older inpatients: the cluster randomized CHERISH trial. JAMA Intern Med 2022;182(3):274-282 [FREE Full text] [doi: 10.1001/jamainternmed.2021.7556] [Medline: 35006265]
18. Daniels H, Hollinghurst J, Fry R, Clegg A, Hillcoat-Nallétamby S, Nikolova S, et al. The value of routinely collected data in evaluating home assessment and modification interventions to prevent falls in older people: systematic literature review. JMIR Aging 2021;4(2):e24728 [FREE Full text] [doi: 10.2196/24728] [Medline: 33890864]
19. Han E, Kharrazi H, Shi L. Identifying predictors of nursing home admission by using electronic health records and administrative data: scoping review. JMIR Aging 2023;6:e42437 [FREE Full text] [doi: 10.2196/42437] [Medline: 37990815]
20. Yasunaga H. Real world data in Japan: chapter II the diagnosis procedure combination database. ACE 2019;1(3):76-79. [doi: 10.37737/ace.1.3_76]
21. Mitsutake S, Ishizaki T, Yano S, Tsuchiya-Ito R, Uda K, Toba K, et al. All-cause readmission or potentially avoidable readmission: which is more predictable using frailty, comorbidities, and ADL? Innov Aging 2023;7(5):igad043 [FREE Full text] [doi: 10.1093/geroni/igad043] [Medline: 37342490]
22. Uda K, Matsui H, Fushimi K, Yasunaga H. Intensive in-hospital rehabilitation after hip fracture surgery and activities of daily living in patients with dementia: retrospective analysis of a nationwide inpatient database. Arch Phys Med Rehabil 2019;100(12):2301-2307. [doi: 10.1016/j.apmr.2019.06.019] [Medline: 31421098]
23. Health survey on impact assessment of DPC introduction: summary of statistics. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/stf/shingi2/0000196043_00006.html; [accessed 2024-10-02]
24. Hospital-acquired complications (HACs). Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/our-work/indicators-measurement-and-reporting/hospital-acquired-complications-hacs [accessed 2024-10-31]
25. Department of health and human services. Patient safety indicators (PSI) benchmark data tables. Agency for Healthcare Research and Quality, U.S. URL: https://qualityindicators.ahrq.gov/Downloads/Modules/PSI/V2024/Version_2024_Benchmark_Tables_PSI.pdf [accessed 2024-10-31]

26. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol 2017;27(10):476-482 [FREE Full text] [doi: 10.1016/j.je.2016.09.009] [Medline: 28142051]
27. Ibitoye T, So S, Shenkin S, Anand A, Reed MJ, Vardy ERLC, et al. Delirium is under-reported in discharge summaries and in hospital administrative systems: a systematic review. Delirium (Bielef) 2023;2023:74541. [doi: 10.56392/001c.74541] [Medline: 39654697]
28. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965;14:61-65. [doi: 10.1037/t02366-000]
29. Exchange rates. Organization for economic co-operation and development. URL: https://www.oecd.org/en/data/indicators/exchange-rates.html [accessed 2024-03-18]
30. Mitsutake S, Ishizaki T, Tsuchiya-Ito R, Furuta K, Hatakeyama A, Sugiyama M, et al. Association of cognitive impairment severity with potentially avoidable readmissions: a retrospective cohort study of 8897 older patients. Alzheimers Dement (Amst) 2021;13(1):e12147 [FREE Full text] [doi: 10.1002/dad2.12147] [Medline: 33816752]
31. Mitsutake S, Ishizaki T, Tsuchiya-Ito R, Furuta K, Hatakeyama A, Sugiyama M, et al. Association of cognitive impairment severity with potentially avoidable readmissions: a retrospective cohort study of 8897 older patients. Alzheimers Dement (Amst) 2021;13(1):e12147 [FREE Full text] [doi: 10.1002/dad2.12147] [Medline: 33816752]
32. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
33. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, et al. Development and validation of a hospital frailty risk score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018;391(10132):1775-1782 [FREE Full text] [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]
34. Tran NN, Hoang TPN, Ho TKT. Diagnosis and risk factors for delirium in elderly patients in the emergency rooms and intensive care unit of the national geriatric hospital emergency department: a cross-sectional observational study. Int J Gen Med 2021;14:6505-6515 [FREE Full text] [doi: 10.2147/IJGM.S325365] [Medline: 34675618]
35. Raats JW, Steunenberg SL, de Lange DC, van der Laan L. Risk factors of post-operative delirium after elective vascular surgery in the elderly: a systematic review. Int J Surg 2016;35:1-6 [FREE Full text] [doi: 10.1016/j.ijsu.2016.09.001] [Medline: 27613124]
36. Hori K, Usuba K, Sakuyama A, Adachi Y, Hirakawa K, Nakayama A, et al. Hospitalization-associated disability after cardiac surgery in elderly patients - exploring the risk factors using machine learning algorithms. Circ Rep 2021;3(8):423-430 [FREE Full text] [doi: 10.1253/circrep.CR-21-0057] [Medline: 34414331]
37. Bates DW, Pruess K, Souney P, Platt R. Serious falls in hospitalized patients: correlates and resource utilization. Am J Med 1995;99(2):137-143. [doi: 10.1016/s0002-9343(99)80133-8] [Medline: 7625418]
38. Moghadamyeghaneh Z, Stamos MJ, Stewart L. Patient co-morbidity and functional status influence the occurrence of hospital acquired conditions more strongly than hospital factors. J Gastrointest Surg 2019;23(1):163-172. [doi: 10.1007/s11605-018-3957-9] [Medline: 30225796]
39. Campbell J, Hubbard R, Ostaszkiewicz J, Green T, Coyer F, Mudge A. Incontinence during and following hospitalisation: a prospective study of prevalence, incidence and association with clinical outcomes. Age Ageing 2023;52(9):afad181 [FREE Full text] [doi: 10.1093/ageing/afad181] [Medline: 37738169]
40. Montero-Marco J, Charlo-Bernardos M, Subirón-Valera AB, Erickson H, Herrero-Cortina B, Altarribas-Bolsa E. The role of nursing care continuity report in predicting length of hospital stay in older people: a retrospective cohort study. J Clin Nurs 2024;33(5):1830-1838. [doi: 10.1111/jocn.16953] [Medline: 38178555]
41. Brand CA, Sundararajan V. A 10-year cohort study of the burden and risk of in-hospital falls and fractures using routinely collected hospital data. Qual Saf Health Care 2010;19(6):e51. [doi: 10.1136/qshc.2009.038273] [Medline: 20558479]
42. Rubin EB, Buehler AE, Halpern SD. States worse than death among hospitalized patients with serious illnesses. JAMA Intern Med 2016;176(10):1557-1559 [FREE Full text] [doi: 10.1001/jamainternmed.2016.4362] [Medline: 27479808]
43. Cordoba G, Schwartz L, Woloshin S, Bae H, Gøtzsche PC. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. Br Med J 2010;341:c3920 [FREE Full text] [doi: 10.1136/bmj.c3920] [Medline: 20719825]
44. Sundararajan V, Quan H, Halfon P, Fushimi K, Luthi JC, Burnand B, International Methodology Consortium for Coded Health Information (IMECCHI). Cross-national comparative performance of three versions of the ICD-10 Charlson index. Med Care 2007;45(12):1210-1215. [doi: 10.1097/MLR.0b013e3181484347] [Medline: 18007172]

Abbreviations

ADL: activities of daily living
AOR: adjusted odds ratio
ARR: adjusted risk ratio
BI: Barthel Index
CCI: Charlson Comorbidity Index

DPC: Diagnosis Procedure Combination
ICD-10: International Statistical Classification of Diseases, Tenth Revision
HAC-OP: hospital-associated complications of older people
HAC-OP-DPC: hospital-associated complications of older people-Diagnosis Procedure Combination data version
HFRS: Hospital Frailty Risk Score
LOS: length of stay
LTCF: long-term care facility

Edited by M Gray; submitted 01.11.24; peer-reviewed by T Karen; comments to author 01.12.24; revised version received 14.12.24; accepted 02.01.25; published 06.02.25.

Please cite as:
Mitsutake S, Ishizaki T, Yano S, Hirata T, Ito K, Furuta K, Shimazaki Y, Ito H, Mudge A, Toba K
Predictive Validity of Hospital-Associated Complications of Older People Identified Using Diagnosis Procedure Combination Data From an Acute Care Hospital in Japan: Observational Study
JMIR Aging 2025;8:e68267
URL: https://aging.jmir.org/2025/1/e68267
doi:10.2196/68267
PMID:

Detecting Sleep/Wake Rhythm Disruption Related to Cognition in Older Adults With and Without Mild Cognitive Impairment Using the myRhythmWatch Platform: Feasibility and Correlation Study

Caleb D Jones[1], MS; Rachel Wasilko[2], MPH; Gehui Zhang[3], PhD; Katie L Stone[4,5], MA, PhD; Swathi Gujral[2], PhD; Juleen Rodakowski[6], MS, OTD, OTR/L; Stephen F Smagula[2], PhD

[1]Primary Care Accelerated Track, School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Psychiatry, School of Medicine, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh, PA, United States

[3]School of Sciences, Southwest Petroleum University, Chengdu, Sichuan, China

[4]California Pacific Medical Center Research Institute, San Francisco, CA, United States

[5]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, United States

[6]Department of Occupational Therapy, School of Health and Rehabilitation Sciences, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Stephen F Smagula, PhD
Department of Psychiatry, School of Medicine, University of Pittsburgh, 3811 O'Hara Street, Pittsburgh, PA, United States

Abstract

Background: Consumer wearable devices could, in theory, provide sufficient accelerometer data for measuring the 24-hour sleep/wake risk factors for dementia that have been identified in prior research. To our knowledge, no prior study in older adults has demonstrated the feasibility and acceptability of accessing sufficient consumer wearable accelerometer data to compute 24-hour sleep/wake rhythm measures.

Objective: We aimed to establish the feasibility of characterizing 24-hour sleep/wake rhythm measures using accelerometer data gathered from the Apple Watch in older adults with and without mild cognitive impairment (MCI), and to examine correlations of these sleep/wake rhythm measures with neuropsychological test performance.

Methods: Of the 40 adults enrolled (mean [SD] age 67.2 [8.4] years; 72.5% female), 19 had MCI and 21 had no cognitive disorder (NCD). Participants were provided devices, oriented to the study software (myRhythmWatch or myRW), and asked to use the system for a week. The primary feasibility outcome was whether participants collected enough data to assess 24-hour sleep/wake rhythm measures (ie, ≥3 valid continuous days). We extracted standard nonparametric and extended-cosine based sleep/wake rhythm metrics. Neuropsychological tests gauged immediate and delayed memory (Hopkins Verbal Learning Test) as well as processing speed and set-shifting (Oral Trails Parts A and B).

Results: All participants meet the primary feasibility outcome of providing sufficient data (≥3 valid days) for sleep/wake rhythm measures. The mean (SD) recording length was somewhat shorter in the MCI group at 6.6 (1.2) days compared with the NCD group at 7.2 (0.6) days. Later activity onset times were associated with worse delayed memory performance (β=−.28). More fragmented rhythms were associated with worse processing speed (β=.40).

Conclusions: Using the Apple Watch-based myRW system to gather raw accelerometer data is feasible in older adults with and without MCI. Sleep/wake rhythms variables generated from this system correlated with cognitive function, suggesting future studies can use this approach to evaluate novel, scalable, risk factor characterization and targeted therapy approaches.

(JMIR Aging 2025;8:e67294) doi:10.2196/67294

KEYWORDS

sleep; sleep/wake; circadian; activity pattern; dementia; cognition; mobile sensing; actigraphy; accelerometer

Introduction

Twenty-four-hour sleep/wake characteristics, which are objectively measurable using accelerometer-containing devices, are related to both dementia biomarkers and dementia risk. Prior studies have shown that sleep/wake rhythm disruption, including fragmentation of 24-hour sleep/wake rhythms, temporally precedes the incidence of mild cognitive impairment (MCI) and dementia [1,2]. Even among adults with normal cognition, rhythm fragmentation correlates with greater brain amyloid deposition [3,4]. Over time, 24-hour rhythm fragmentation has been associated with increased risk and faster rates of cognitive decline in people with MCI and mild-to-moderate dementia [1,5] .

The above-mentioned studies linking sleep/wake rhythm disruption with cognitive impairment and neurodegenerative processes in aging raise the potential that accelerometer-based sleep/wake monitoring may have useful clinical applications. For example, wearable accelerometers could potentially be used to identify individuals who have established sleep/wake risk factors for dementia, assign targeted interventions, and track sleep/wake patterns throughout clinical trials. Compared with researcher-focused accelerometer devices, consumer wearables (which also contain triaxial accelerometers) could yield more widely scalable, and clinician and user-friendly, systems for developing and testing potential applications. Our prior pilot study demonstrated that it is possible to collect 24-hour accelerometer data from the Apple Watch and generate standard 24-hour sleep/wake rhythm measures [6]. However, this prior study was limited to a convenience sample of young adults.

Prior to studies evaluating the use of consumer wearable-based sleep/wake monitoring for dementia risk stratification and targeting prevention approaches, we sought to establish the feasibility of using a consumer wearable device to assess 24-hour sleep/wake patterns in older adults (including those with elevated dementia risk by virtue of having a diagnosis of MCI). Our first aim was therefore to evaluate the feasibility of using the Apple Watch and a software platform called myRhythmWatch (myRW) to obtain 24-hour accelerometer data assessing sleep/wake rhythms in older adults with and without MCI. Second, we sought to validate that the sleep/wake rhythm data collected from this system is relevant to cognition. To do so, we evaluated if sleep/wake rhythm variables extracted from this system correlated with cognitive function similar to prior published studies that used researcher-focused accelerometer devices.

Methods

Participants and Study Protocol

Participants were identified by referral from either local studies that adjudicated MCI diagnoses or a local research recruitment registry that is led by the University of Pittsburgh Clinical and Translational Science Institute (Pitt+Me). The inclusion criteria were (1) being 50 years of age or older; (2) passing the San Diego Brief Assessment of Capacity to Consent [7] with scores ≥14.5; (3) have a score of >27 on the Telephone Interview for Cognitive Status (TICS) [8]; (4) having a TICS score of either ≤34 (high-risk group) or ≥39 (no cognitive disorder or NCD); and (5) have a prior adjudicated diagnosis of MCI (high-risk group) or reporting no concerns regarding cognitive decline (ie, NCD). The exclusion criteria included (1) self-reporting active behavioral health treatments for insomnia or depression; (2) self-reporting of performing the prescribed exercises; and (3) self-reported use of sleep medications every night or nearly every night. After explaining the study purpose and procedures, we obtained verbal consent to screen interested individuals for eligibility. Of 82 potentially eligible participants who completed the verbal eligibility screening, 35 were ineligible, 6 refused, and 41 were enrolled. One participant withdrew after enrollment, resulting in an analytic sample of 40 individuals (19 high-risk (ie, having a low TICS score and prior MCI diagnosis) and 21 low-risk (ie, having a high TICS score and no prior MCI diagnosis), as shown in Figure S1 in Multimedia Appendix 1 (ie, a flow diagram illustrating how we arrived at our analytic sample).

After completing web-based written informed consent forms, participants completed baseline procedures and assessments via Health Insurance Portability and Accountability Act (HIPPA)-compliant video-conferencing. We sent participants the following study devices: an Apple Watch 8, iPhone SE 2nd generation with active data plans, and nonstock 40-Watt charger. After the participants received the devices, we conducted an additional study visit to instruct them on logging in and using the myRW application. We specifically instructed participants that about 20 - 30 minutes of charging a day is sufficient, and asked them to wear the watch whenever it was not charging for a week.

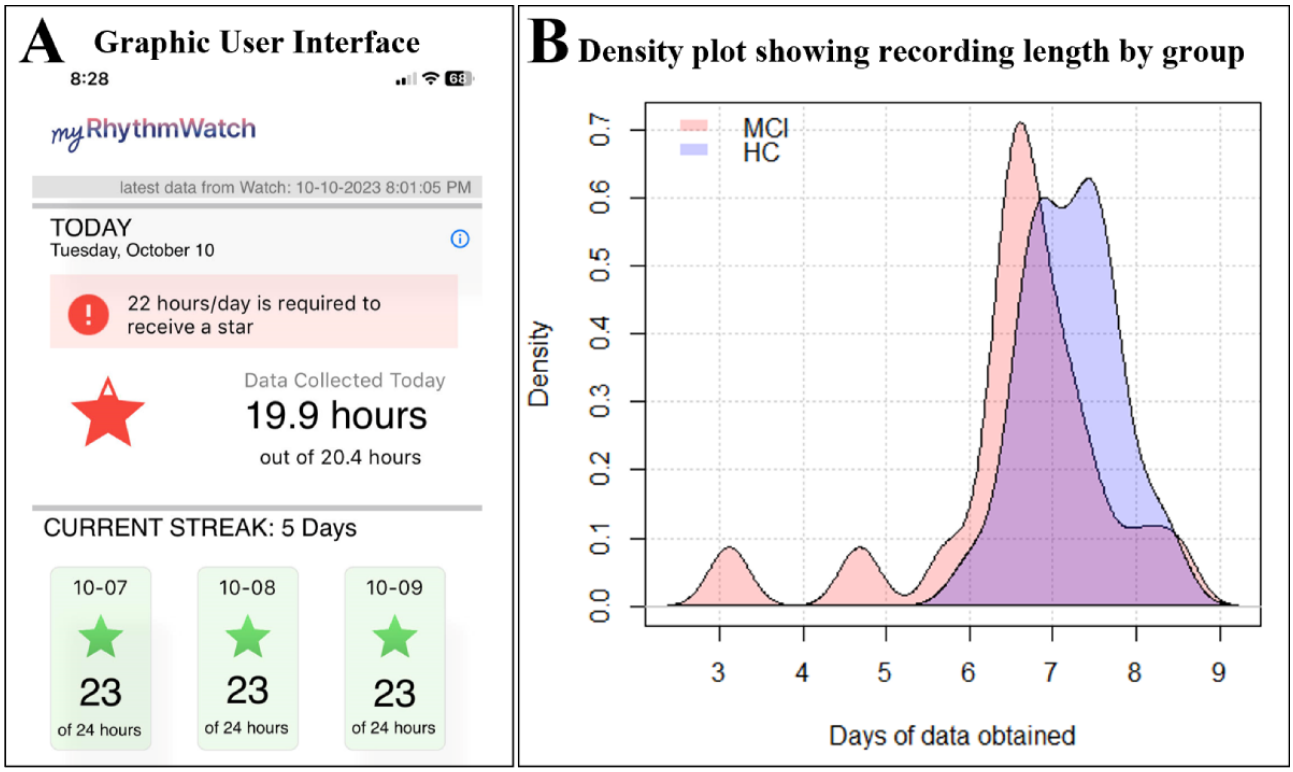
Ethical Considerations

All study procedures were approved by the University of Pittsburgh Institutional Review Board (identifier: STUDY22080033). All the participants completed web-based written informed consent forms

Study Application

The myRW software system, developed by the corresponding author, extracts triaxial accelerometer data that are recorded at 50 hertz on the Apple Watch. After bandpass filtering, we aggregate across axes and into 30-second counts, and transmit data for storage and further computations on the Amazon Web Services platform. To encourage data collection, we depict the amount of data collected per day graphically by proportionally filling and color coding a star (Figure 1). We explicitly told participants that sufficient data were required to “get a star,” that the goal was to get 7 stars in a row, and we also displayed information on how many days in a row participants obtained stars (to encourage data collection “streaks”). Participants received reminders three times daily to check in to the application. Other features, including personal data graphs and sleep/wake metrics, were disabled for this study to prevent the feedback from potentially altering the users’ behavior/typical sleep/wake patterns.

Figure 1. Illustrations of the graphic user interface (left) and the amount of accelerometer data collected in the sample (right). The difference in the mean (SD) recording length between the no cognitive disorder (NCD; 7.2 [0.6]) and mild cognitive impairment (MCI; 6.6 [1.2]) groups is shown visually in the density plot (*P*=.04).



Sleep/Wake Measures

Consistent with prior publications using researcher-focused accelerometers [9-11], we required participants to have at least 3 valid continuous days of accelerometer data to consider the recording adequate for processing. Valid days were defined, consistent with National Sleep Research Resource data processing standards [12], as those with no more than 4 hours of nonwear time or any nonwear time during the main sleep period. Invalid days and remaining data missing within days were censored (not imputed). Although we applied this aforementioned 20 hours per day criterion, to maximize data collection, note that the app was programmed to reward participants with a green star only if they collected at least 22 hours of data per day.

Following previously published technical definitions, we calculated extended cosine-based [13] and nonparametric [14-17] sleep/wake measures using the R package 'RAR' and custom code. From the extended-cosine models, we extracted measures of 24-hour robustness (pseudo-F statistic, indicating how well the observed data fits the 24-hour curve); activity onset time (up-mesor, the time which the modeled activity level passes the middle modeled rhythm height prior to the peak); and activity offset time (down-mesor or the time which the modeled activity level passes the middle modeled rhythm height prior to the nadir). From the nonparametric method, we calculated the cross-daily stability (inter-daily stability, measuring the consistency of circadian sleep/wake activity rhythms across days); rhythm strength (relative amplitude, measuring the standardized peak-trough difference of 24-hour activity rhythms); and sleep/wake rhythm fragmentation (intradaily variability or IV, measuring the frequency and extent of transitions in activity levels). Although sleep/wake rhythm fragmentation is typically measured using differences in activity levels every hour, it can also be computed on a range of timescales [16,17]. For all nonparametric metrics, we used all-time series data and did not subsample to hourly activity levels, as described previously [14]. To explore the relevance of fragmentation timescale, we computed IV values using numerator timescales ranging from 5 to 60 minutes, but holding the denominator constant as done previously [16].

Neuropsychological Measures

Cognitive performance was measured using the Hopkins Verbal Learning Test (HVLT; PAR, Lutz, FL) and the Oral Trail Making Test (O-TMT).

The HVLT is a verbal list learning measure with three trials of 12 words from multiple representative semantic categories, and includes delayed recall and delayed recognition memory components [18,19]. For the delayed recall, participants were asked to repeat as many terms as they could recall from the word listed presented to them thrice, 30 minutes earlier. For delayed memory recognition, participants were asked to recognize those same terms appearing on a broader list containing semantically related terms, ie, "horse" must be differentiated from "dog," which was not part of the original term set. The HVLT assessed encoding, storage, and retrieval of noncontextual verbal information.

Processing speed was measured as the time taken on Part A of the O-TMT, which requires the subject to verbally count upwards from 1 to 25 as quickly as they are able [20]. Set-shifting, an aspect of cognitive function, was measured as

the time taken using Part B of the O-TMT, which requires the subject to alternate between listing numbers and letters in sequence (“1, A, 2, B,” etc). Note that longer times on both parts of the O-TMT indicates worse performance [21].

Statistical Analyses

To compare recording lengths between the MCI and NCD groups, we used an independent sample *t* test and density plots illustrating recording length distributions by groups. To examine relationships of sleep/wake rhythm measures with cognitive function, we first adjusted for key confounders by taking the residuals from linear regression models (one for each cognitive outcome variable) that had the confounders age, sex, education (college degree vs less than college degree), and accelerometer recording length as predictor variables. We used the residualized values as age, sex, education, and recording length-adjusted cognitive outcome variables in a series of linear regressions (one per sleep/wake predictor variable). Since there were many conceptually similar or highly inter-correlated intradaily variability metrics, for related analyses, we only report Benjamini-Hochberg corrected false discovery rates (which are here referred to as *q* values instead of *P* values) [22] to account for the 12 statistical tests relating intradaily variability metrics within each cognitive outcome.

Results

Sample Characteristics

The sample included older adults, with a mean (SD) age of 67.2 (8.4) years; the majority of the participants were females with college degrees (Table 1).

Table . Sample characteristics.

Characteristics	Value
Age, years, mean (SD)	67.2 (8.4)
Female sex, n (%)	29 (72.5)
College degree or greater, n (%)	29 (72.5)
Prior diagnosis of MCI[a], n (%)	19 (47.5)
Accelerometer recording length, days, mean (SD)	6.9 (0.9)
Immediate memory, mean (SD)	26.7 (5.2)
Delayed memory, mean (SD)	8.7 (2.8)
Processing speed, seconds, mean (SD)	10.3 (2.8)
Set shifting executive function, seconds, mean (SD)	40.1 (31.6)

[a]MCI: mild cognitive impairment

Accelerometer Recording Lengths in People With and Without a Diagnosis of MCI

In both the groups, all participants achieved the minimum data requirement for computing 24-hour sleep/wake rhythm assessments. There was, however, a statistically significant difference in the mean (SD) recording lengths of about a half a day when comparing the NCD (7.2 [0.6] days) and MCI groups (6.6 [1.2] days), with the Satterthwaite test assuming unequal group variances (*df* 25.033; *t* value −2.15; *P*=.04). This difference was due to 3 individuals in the MCI group who collected 3 - 6 days of data each (Figure 1).

Associations of Sleep/Wake Characteristics With Cognitive Performance

As shown in Table 2, there were small effect size associations of later activity onset time with lower delayed memory performance (β=−.28, 95% CI −0.55 to - 0.02; *t*=−2.17, *df*=38, *P*=.04) and more stable rhythms with better processing speed and performance (β=−0.27, 95% CI: −0.54 to 0.00; *t*=−2.00, *df*=38, *P*=.05). None of the other sleep/wake measures listed in Table 2 were associated with the cognitive outcomes. Regarding sleep/wake rhythm fragmentation, greater fragmentation levels in the 40 - 60-minute timeframe was significantly associated with worse processing speed and performance (Table 3; β point estimate range: 0.36, 0.40; *q*=.022). When repeating analyses in the subgroup with at least 6 days of data, results were not substantively altered (n=37; see Tables S1 and S2 in Multimedia Appendix 1).

To illustrate the accelerometer data collected in this study and visualize 24-hour sleep/wake fragmentation, Figure 2 shows data from example participants with lower and higher degrees of 24-hour sleep/wake rhythm fragmentation.

Table . Associations between 24-hour sleep/wake rhythm variables and cognitive performance.

Variable	Immediate memory[a]		Delayed memory[a]		Psychomotor speed/attention[b]		Set shifting executive function[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
24 h robustness	.06 (−.22 to .34)	.67	.17 (−.11 to 0.44)	.22	−.09 (−.38 to .20)	.53	−.18 (−.48 to .11)	.22
Cross-daily stability	.11 (−.17 to .39)	.44	.13 (−.14 to .41)	.34	−.27 (−.54 to .00)	.05	−.11 (−.42 to .19)	.45
Rhythm strength	.08 (−.21 to .36)	.59	.04 (−.23 to .32)	.75	−.04 (−.33 to .24)	.76	−.12 (−.42 to .18)	.43
Activity onset time	−.13 (−.41 to .15)	.36	−.28 (−.55 to -.02)	.04	.04 (−.24 to .33)	.76	−.11 (−.41 to .19)	.46
Activity offset time	−.03 (−.32 to .25)	.81	−.07 (−.35 to .21)	.60	−.12 (−.41 to .16)	.40	.17 (−.13 to .47)	.26

[a]Higher scores on the memory tests indicates better performance (as the outcome is number of items recalled)

[b]For psychomotor speed/attention and set-shifting, higher scores indicate worse performance (as the outcome is the duration of time to complete Oral Trails A and B, respectively)

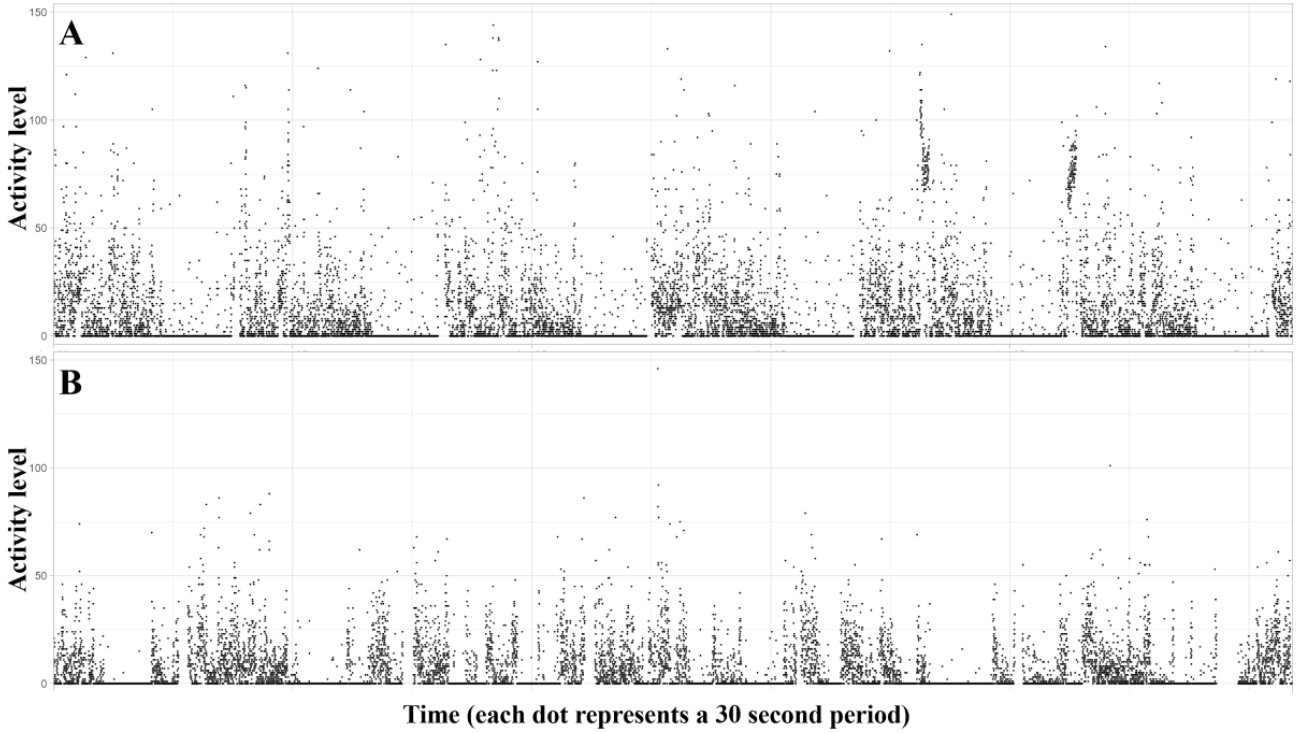
Table . Associations between 24-hour sleep/wake rhythm fragmentation on various timescales with cognitive performance.

Time scale	Immediate memory[a]		Delayed memory[a]		Psychomotor speed/attention[b]		Set shifting executive function[b]	
	β (95% CI)	*q* value	β (95% CI)	*q* value	β (95% CI)	*q* value	β (95% CI)	*q* value
5 minutes	-.10 (−.38 to .18)	0.938	−.01 (−.29 to .27)	0.938	.14 (−.14 to .43)	0.314	.28 (−.01 to .58)	0.092
10 minutes	−.10 (−.38 to .18)	0.938	−.02 (−.3 to .26)	0.938	.14 (−.14 to .43)	0.314	.26 (−.03 to .55)	0.102
15 minutes	−.10 (−.39 to .18)	0.841	−.05 (−.33 to .23)	0.841	.16 (−.12 to .45)	0.298	.30 (.02 to .59)	0.092
20 minutes	−.12 (−.41 to .16)	0.628	−.11 (−.38 to .17)	0.628	.21 (−.07 to .49)	0.203	.34 (.06 to .62)	0.092
25 minutes	−.11 (−.40 to .17)	0.628	-.10 (-.38 to .18)	0.628	.20 (−.08 to .48)	0.203	.34 (.06 to .63)	0.092
30 minutes	−.11 (−.39 to .17)	0.628	−.11 (−.39 to .16)	0.628	.23 (−.05 to .51)	0.177	.32 (.03 to .61)	0.092
35 minutes	−.09 (−.37 to .19)	0.628	−.13 (−.41 to .14)	0.628	.28 (.01 to .55)	0.085	.28 (−.01 to .57)	0.093
40 minutes	−.08 (−.36 to .20)	0.628	−.15 (−.42 to .13)	0.628	.37 (.10 to .63)	0.022	.32 (.03 to .60)	0.092
45 minutes	-.09 (-.38 to .19)	0.628	−.16 (−.44 to .11)	0.628	.37 (.11 to .63)	0.022	.29 (.00 to .58)	0.092
50 minutes	−.11 (−.39 to .17)	0.628	−.18 (−.45 to .09)	0.628	.40 (.14 to .65)	0.022	.25 (−.05 to .54)	0.118
55 minutes	−.08 (−.36 to .21)	0.628	−.18 (−.45 to .1)	0.628	.36 (.09 to .62)	0.022	.17 (−.13 to .47)	0.254
60 minutes	−.12 (−.40 to .16)	0.628	−.23 (−.50 to .04)	0.628	.36 (.09 to .62)	0.022	.18 (−.12 to .48)	0.250

[a]Higher scores on the memory tests indicates better performance (as the outcome is number of items recalled).

[b]For psychomotor speed/attention and set-shifting, higher scores indicate worse performance (as the outcome is duration of time to complete Oral Trails A and B, respectively).

Figure 2. Accelerometer data from two participants. Top: The participant was in the low risk (no cognitive disease) group and had relatively lower sleep/wake rhythm fragmentation. Bottom: The participant is from the high-risk (mild cognitive impairment) group and had relatively higher sleep/wake rhythm fragmentation.



Discussion

These results demonstrate that it is feasible to collect raw accelerometer data and characterize 24-hour sleep/wake rhythms in older adults with and without MCI using the Apple Watch and myRW system. Each individual in our sample met the minimum data requirement to derive sleep/wake rhythms measures. Notably, however, several individuals in the MCI group generated shorter recordings. Thus, while supporting the feasibility of using the Apple Watch and myRW system to assess sleep/wake rhythms in people with MCI who are at risk for dementia, our findings also suggest that some individuals with MCI may require additional support (eg, additional human support or programmed reminders) and data imputation [23] to monitor sleep/wake rhythms over longer periods with this system.

With regard to our second aim, we found that sleep/wake rhythm variables extracted from the Apple Watch accelerometer data were correlated with cognitive function. We found small-to-medium effect size correlations between sleep/wake measures and cognitive performance. Specifically, we found that later activity onset times were related to worse delayed memory performances. In addition, less stable and more fragmented rhythms were correlated with worse processing speed. Detecting these signs of disrupted sleep/wake rhythms early on could help target prevention strategies, given prior research demonstrating that sleep/wake rhythm fragmentation [1], worse memory [24,25], and slower processing speeds [26] are all associated with the risk of developing dementia.

The study had several limitations. This was a cross-sectional observational study limited to accelerometer and neuropsychological measures, so there is no way to determine causality or the mechanisms underlying associations between the variables examined. The sample size was relatively small; therefore, there is a risk of false negative associations and effect size estimates (relating sleep/wake measures with cognition) that should be deemed less reliable than those from large epidemiologic studies. We failed to detect some associations that were expected based on prior literature examining dementia risk (eg, previous findings linking low rhythm strength with dementia risk [27]). Larger studies examining the relationships between these sleep/wake factors and specific domains of cognitive function, earlier in the disease processes, will be needed to verify our findings. We made efforts to minimize data loss as described above, but did not use data imputation, which could be applied in future studies. Additionally, given the small sample, future studies will be required to confirm results of using this system in samples that are more broadly representative, eg, samples including more ethnically diverse population subgroups. Finally, we note that determining the mechanism underlying these relationships between sleep/wake rhythm disruption, cognition, and dementia risk is outside the scope of this work. Recent literature has notably suggested that sleep/wake rhythm fragmentation relates to neurodegeneration of the locus coeruleus [28], which could hasten dementia pathology and cognitive decline [29,30].

In summary, we have demonstrated that it is feasible to use the Apple Watch and myRW system to gather accelerometer data and characterize sleep/wake risk factors for dementia in older adults including adults with MCI. A strength of this study is the use of the Apple Watch, which is already voluntarily being used by millions of people, as it may provide increased scalability for applications of sleep/wake risk factor monitoring into the general public and general practice settings. One implication

of this research is that it is feasible for future research to be conducted for evaluating if monitoring sleep/wake disruption using consumer wearable-based systems improves upon existing dementia risk factor detection and management approaches. Future studies will also be needed to examine if tailoring interventions using information on sleep/wake patterns derived from this system improves outcomes among older adults who are at risk for dementia.

Acknowledgments

CJ was supported by T32MH019986 in completing this work. The work presented herein was supported by a grant from the NIA to SFS (R21AG074094) and to JUR (R01AG056351).

Conflicts of Interest

SFS is owner and chief executive of a start-up, that was a University Licensed start-up, which worked on developing the myRW system with support from an National Institute on Aging STTR award (R41AG069596). KS reports grant funding from Eli Lilly and is supported by grant funding from NIH, PCORI, and DOD.

Multimedia Appendix 1
Supplementary Tables 1 and 2 and Figure 1.
[DOCX File, 36 KB - aging_v8i1e67294_app1.docx]

References

1. Li P, Gao L, Gaba A, et al. Circadian disturbances in Alzheimer's disease progression: a prospective observational cohort study of community-based older adults. Lancet Healthy Longev 2020 Dec;1(3):e96-e105. [doi: 10.1016/s2666-7568(20)30015-5] [Medline: 34179863]
2. Haghayegh S, Gao C, Sugg E, et al. Association of rest-activity rhythm and risk of developing dementia or mild cognitive impairment in the middle-aged and older population: prospective cohort study. JMIR Public Health Surveill 2024 May 7;10:e55211. [doi: 10.2196/55211] [Medline: 38713911]
3. Musiek ES, Bhimasani M, Zangrilli MA, Morris JC, Holtzman DM, Ju YES. Circadian rest-activity pattern changes in aging and preclinical Alzheimer disease. JAMA Neurol 2018 May 1;75(5):582-590. [doi: 10.1001/jamaneurol.2017.4719] [Medline: 29379963]
4. Nguyen Ho PT, Hoepel SJW, Rodriguez-Ayllon M, Luik AI, Vernooij MW, Neitzel J. Sleep, 24-hour activity rhythms, and subsequent amyloid-β pathology. JAMA Neurol 2024 Aug 1;81(8):824-834. [doi: 10.1001/jamaneurol.2024.1755] [Medline: 38913396]
5. Targa ADS, Benítez ID, Dakterzada F, et al. The circadian rest-activity pattern predicts cognitive decline among mild-moderate Alzheimer's disease patients. Alzheimers Res Ther 2021 Sep 25;13(1):161. [doi: 10.1186/s13195-021-00903-7] [Medline: 34563258]
6. Smagula SF, Stahl ST, Krafty RT, Buysse DJ. Initial proof of concept that a consumer wearable can be used for real-time rest-activity rhythm monitoring. Sleep 2022 Mar 14;45(3):zsab288. [doi: 10.1093/sleep/zsab288] [Medline: 34931683]
7. Jeste DV, Palmer BW, Appelbaum PS, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]
8. Ferrucci L, Del Lungo I, Guralnik JM, et al. Is the telephone interview for cognitive status a valid alternative in persons who cannot be evaluated by the mini mental state examination? Aging Clin Exp Res 1998 Aug;10(4):332-338. [doi: 10.1007/BF03339796]
9. Smagula SF, Gasperetti CE, Buysse DJ, et al. Efficacy of the transdiagnostic intervention for sleep and circadian dysfunction for depression symptoms and sleep-wake disruption in older and younger adults: secondary age-stratified analysis of a randomized controlled trial. Am J Geriatr Psychiatry 2024 Apr;32(4):478-488. [doi: 10.1016/j.jagp.2023.11.003] [Medline: 38040569]
10. Smagula SF, Zhang G, Krafty RT, et al. Sleep-wake behaviors associated with cognitive performance in middle-aged participants of the Hispanic Community Health Study/Study of Latinos. Sleep Health 2024 Aug;10(4):500-507. [doi: 10.1016/j.sleh.2024.02.002] [Medline: 38693044]
11. Lau SCL, Zhang G, Rueschman M, et al. Sleep-wake behavioral characteristics associated with depression symptoms: findings from the multi-ethnic study of atherosclerosis. Sleep 2024 Apr 12;47(4):zsae045. [doi: 10.1093/sleep/zsae045] [Medline: 38394355]
12. Dean DA 2nd, Goldberger AL, Mueller R, et al. Scaling up scientific discovery in sleep medicine: the national sleep research resource. Sleep 2016 May 1;39(5):1151-1164. [doi: 10.5665/sleep.5774] [Medline: 27070134]
13. Marler MR, Gehrman P, Martin JL, Ancoli-Israel S. The sigmoidally transformed cosine curve: a mathematical model for circadian rhythms with symmetric non-sinusoidal shapes. Stat Med 2006 Nov 30;25(22):3893-3904. [doi: 10.1002/sim.2466] [Medline: 16381069]

14. Suibkitwanchai K, Sykulski AM, Perez Algorta G, Waller D, Walshe C. Nonparametric time series summary statistics for high-frequency accelerometry data from individuals with advanced dementia. PLoS ONE 2020;15(9):e0239368. [doi: 10.1371/journal.pone.0239368] [Medline: 32976498]
15. Gonçalves BSB, Adamowicz T, Louzada FM, Moreno CR, Araujo JF. A fresh look at the use of nonparametric analysis in actimetry. Sleep Med Rev 2015 Apr;20(84-91):84-91. [doi: 10.1016/j.smrv.2014.06.002] [Medline: 25065908]
16. Fossion R, Rivera AL, Toledo-Roy JC, Ellis J, Angelova M. Multiscale adaptive analysis of circadian rhythms and intradaily variability: application to actigraphy time series in acute insomnia subjects. PLoS ONE 2017;12(7):e0181762. [doi: 10.1371/journal.pone.0181762] [Medline: 28753669]
17. Gonçalves BSB, Cavalcanti PRA, Tavares GR, Campos TF, Araujo JF. Nonparametric methods in actigraphy: an update. Sleep Sci 2014 Sep;7(3):158-164. [doi: 10.1016/j.slsci.2014.09.013] [Medline: 26483921]
18. Vanderploeg RD, Schinka JA, Jones T, Small BJ, Graves AB, Mortimer JA. Elderly norms for the Hopkins verbal learning test-revised. Clin Neuropsychol 2000 Aug;14(3):318-324. [doi: 10.1076/1385-4046(200008)14:3;1-P;FT318] [Medline: 11262707]
19. Tri Budi Rahardjo XX. The Hopkins verbal learning test and detection of MCI and mild dementia: a literature review. J Alzheimers Dis Parkinsonism 2014 Oct;04(6):166. [doi: 10.4172/2161-0460.1000166]
20. Mrazik M, Millis S, Drane DL. The oral trail making test: effects of age and concurrent validity. Arch Clin Neuropsychol 2010 May;25(3):236-243. [doi: 10.1093/arclin/acq006] [Medline: 20197294]
21. Ashendorf L, Jefferson AL, O'Connor MK, Chaisson C, Green RC, Stern RA. Trail making test errors in normal aging, mild cognitive impairment, and dementia. Arch Clin Neuropsychol 2008 Mar;23(2):129-137. [doi: 10.1016/j.acn.2007.11.005] [Medline: 18178372]
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
23. Weed L, Lok R, Chawra D, Zeitzer J. The impact of missing data and imputation methods on the analysis of 24-hour activity patterns. Clocks Sleep 2022 Sep 27;4(4):497-507. [doi: 10.3390/clockssleep4040039] [Medline: 36278532]
24. Hayat SA, Luben R, Khaw KT, Brayne C. The relationship between cognitive performance using tests assessing a range of cognitive domains and future dementia diagnosis in a British cohort: a ten-year prospective study. J Alzheimers Dis 2021;81(1):123-135. [doi: 10.3233/JAD-210030] [Medline: 33867360]
25. Rajan KB, Wilson RS, Weuve J, Barnes LL, Evans DA. Cognitive impairment 18 years before clinical diagnosis of Alzheimer disease dementia. Neurology (ECronicon) 2015 Sep 8;85(10):898-904. [doi: 10.1212/WNL.0000000000001774] [Medline: 26109713]
26. Amieva H, Meillon C, Proust-Lima C, Dartigues JF. Is low psychomotor speed a marker of brain vulnerability in late life? Digit symbol substitution test in the prediction of Alzheimer, Parkinson, stroke, disability, and depression. Dement Geriatr Cogn Disord 2019;47(4-6):297-305. [doi: 10.1159/000500597] [Medline: 31466055]
27. Winer JR, Lok R, Weed L, et al. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. Alzheimers Res Ther 2024 Feb 14;16(1):35. [doi: 10.1186/s13195-024-01411-0] [Medline: 38355598]
28. Van Egroo M, van Someren EJW, Grinberg LT, Bennett DA, Jacobs HIL. Associations of 24-hour rest-activity rhythm fragmentation, cognitive decline, and postmortem locus coeruleus hypopigmentation in Alzheimer's disease. Ann Neurol 2024 Apr;95(4):653-664. [doi: 10.1002/ana.26880] [Medline: 38407546]
29. Chen Y, Chen T, Hou R. Locus coeruleus in the pathogenesis of Alzheimer's disease: a systematic review. Alzheimers Dement (N Y) 2022;8(1):e12257. [doi: 10.1002/trc2.12257] [Medline: 35282658]
30. Bueichekú E, Diez I, Kim CM, et al. Spatiotemporal patterns of locus coeruleus integrity predict cortical tau and cognition. Nat Aging 2024 May;4(5):625-637. [doi: 10.1038/s43587-024-00626-y] [Medline: 38664576]

Abbreviations

HVLT: Hopkins Verbal Learning Test
MCI: mild cognitive impairment
NCD: no cognitive disorder
O-TMT: Oral Trail Making Test

Edited by AF Abdulai; submitted 21.10.24; peer-reviewed by R Lok, S Miyata; revised version received 04.03.25; accepted 07.03.25; published 07.04.25.

Please cite as:
Jones CD, Wasilko R, Zhang G, Stone KL, Gujral S, Rodakowski J, Smagula SF
Detecting Sleep/Wake Rhythm Disruption Related to Cognition in Older Adults With and Without Mild Cognitive Impairment Using the myRhythmWatch Platform: Feasibility and Correlation Study
JMIR Aging 2025;8:e67294
URL: https://aging.jmir.org/2025/1/e67294
doi:10.2196/67294

Identifying Optimal Wearable Devices for Monitoring Mobility in Hospitalized Older Adults: Feasibility, Acceptability, and Validity Study

Paulo Nascimento[1], PhD; Renata Kirkwood[1], PhD; Lauren E Griffith[2], PhD; Mylinh Duong[3,4], PhD; Cody Cooper[1], BEng; Yujiao Hao[5], PhD; Rong Zheng[5], PhD; Samir Raza[4], MD; Marla Beauchamp[1], PhD

[1]School of Rehabilitation Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

[2]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

[3]Population Health Research Institute, McMaster University, Hamilton, ON, Canada

[4]Department of Medicine, McMaster University, Hamilton, ON, Canada

[5]Department of Computing and Software, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Marla Beauchamp, PhD
School of Rehabilitation Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Hospitalized, frail older adults have an increased risk of developing hospital-acquired disability associated with hospital practices of restricted physical activity and immobilization. The use of activity tracking wearable devices may allow identification and prevention of mobility decline, reducing hospital-acquired disability.

Objective: This study aimed to identify the optimal wearable device and wear location for monitoring mobility in older hospitalized patients. Specific objectives included (1) comparison of the feasibility and acceptability of ActiGraph wGT3X-BT (ActiGraph LLC), MOX1 (Maastricht Instruments), MetaMotionC (mBientLab), and Fitbit Versa (Google) for continuous mobility monitoring and (2) determination of the concurrent validity of the selected device for detecting body posture and step count.

Methods: Participants were recruited for this observational study in the acute medical care unit of an academic hospital in Hamilton, Ontario, Canada. Eligible patients were aged 60 years and older, able to undertake the mobility protocol, and had an anticipated length of stay greater than 4 days. The study was divided into 2 experiments. Experiment 1 evaluated the feasibility of 4 wearable devices and validated the derived data for body posture and step count. Experiment 2 involved a mobility assessment session and a 24-hour monitoring and feasibility period with the selected device from experiment 1.

Results: The ActiGraph wGT3X-BT emerged as the most feasible device, demonstrating superior usability, data acquisition, and management. The thigh-worn ActiGraph accurately detected sedentary behavior, while the ankle-worn device provided detailed information on step counts and body postures. Bland-Altman plots and intraclass correlation coefficients indicated that the ankle-worn ActiGraph showed excellent reliability for step counting, with minimal bias and narrow limits of agreement. Patients expressed a high willingness to wear a continuous mobility tracking device at the hospital and at home.

Conclusions: Thigh- and ankle-worn ActiGraph are optimal for assessing and monitoring mobility in older hospitalized patients. Challenges such as discomfort and device removal observed during the 24-hour monitoring period highlight areas for future studies. Overall, our findings support the integration of wearable technology in hospital settings to enhance mobility monitoring and early intervention strategies. Further research is warranted to evaluate the long-term use of wearable data for predicting health outcomes post hospitalization and informing clinical decision-making to promote early mobility.

(JMIR Aging 2025;8:e64372) doi:10.2196/64372

KEYWORDS

older adults; gerontology; geriatric; aging; feasibility; acceptability; mobility; wearable; inpatient; hospital-acquired disability; physical performance; mHealth; mobile health; hospital; physical activity; exercise; Fitbit; posture; walk

Introduction

Research indicates that a significant proportion of the adverse functional and health outcomes experienced by older adults during hospitalization may not be directly linked to their underlying health conditions or the reasons for hospitalization [1-3]. Instead, they may arise from certain hospital practices, such as restricted physical activity and immobilization, which may be harmful for older patients, especially for those who are frail [1]. Studies have consistently shown that older patients

spend a minimal amount of time standing or walking during their hospital stay, typically just 3% of the time [4-7]. This limited mobility can have significant consequences for function, with each day spent in bed associated with a 1% - 5% loss in muscle strength [3]. On the other hand, there is evidence showing that older adults who regain their prehospitalization level of function after discharge have lower mortality rates and maintain their functional levels 1 year post discharge. Therefore, early detection and prevention of mobility decline during hospitalization is critical to improving patient outcomes and reducing health care utilization [8].

Wearable technology provides a direct means of assessing and monitoring mobility, by gathering continuous information on patients' physical activities and mobility patterns. This capability would allow practitioners to create and monitor tailored mobility care for each patient, which may improve functional outcomes. Several wearable monitors have been validated in healthy individuals including among older adults living in the community [9,10]. However, experiment protocols, including population type, settings, sensor type, activity, and wear location (eg, wrist vs thigh) can influence the validity of these devices [11-14]. Despite the increased use of wearables in clinical settings, the feasibility, validity, and reliability of these devices have not been fully established in older hospitalized patients [15,16]. Collecting data accurately with wearables in hospitalized patients presents significant challenges, as they tend to be more sedentary and walk slower [17]. In addition, hospitalized patients often have medical devices attached to them, such as intravenous lines or heart monitors, which can impede mobility and make it challenging to select the optimal wear location for data collection.

The overall objective of this study was to identify the optimal wearable device and wear location for assessing and monitoring mobility in older hospitalized patients. Specifically, we aimed to (1) compare the feasibility and acceptability of different wearable devices to assess mobility for long-term continuous monitoring during inpatient hospital stays among older hospitalized patients, and (2) for the selected device from aim 1, determine its concurrent validity for detecting body posture and step count. To fulfill the study aims we performed 2 experiments with 2 independent samples. In the first experiment, we used a standardized protocol to test the feasibility of data collection with the ActiGraph wGT3X-BT, MOX 1, MetaMotionC (MMC), and Fitbit Versa over a short period of time, while validating the data collected for detecting body posture and step count using a standardized protocol. The second experiment was divided in 2 parts where patients wore the selected wearable from experiment 1 in 3 different body locations and performed a standardized activity protocol, followed by a 24-hour free-living protocol. This investigation was crucial for understanding the practicality of wearable devices for mobility monitoring in older adults in a hospital setting, as well as identifying potential barriers or limitations that may impact their utility.

Methods

Participants

Patients were recruited from the acute medical care unit of the Juravinski Hospital, Hamilton, Ontario, Canada. Inpatients were initially identified by the admitting physician, who sought approval before study personnel approached them. Patients were only included if, based on the judgment of the admitting physician, they were deemed capable of providing informed consent. A study coordinator, not directly involved in patient care, then contacted the eligible patients to obtain informed consent. During data collection, a hospital physical therapist, who accompanied all participants and assisted with data collection, ensured safety and facilitated communication when necessary.

Eligible patients were aged 60 years and older, able to undertake the mobility protocol with or without assistance, had an anticipated length of stay of more than 4 days, and able to provide written informed consent. Experiment 1 was conducted from November to December of 2019 and experiment 2 from September to December 2020.

Ethical Considerations

This study was approved by the Hamilton Integrated Research Ethics Board (HiREB #7145) and all participants provided written informed consent before participation. At the time of consent, we recorded information regarding participants' demographics, preadmission functional performance, and health status. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants did not receive compensation for their participation.

Measurement Instruments

For the feasibility and device performance of experiment 1, 4 devices were compared, that is, the MetaMotionC (MMC), Fitbit Versa 1, MOX1, and the medical-grade monitor ActiGraph wGT3X-BT. The MMC by mBientLab was chosen for its open platform enabling on-board programming, in addition to its lower cost. The Fitbit Versa 1 (Google) was widely recognized in the community at the study's time, despite lacking direct access to raw sensor data extraction. For both devices, a Python algorithm was developed to extract the raw data. The MOX1 device (Maastricht Instruments , Netherlands) presents a triaxial accelerometer sensor and is equipped with proprietary software. During the study period, the proprietary software of the MOX1 did not offer direct access to the raw data and step count data. To address this limitation, the company provided us with a MATLAB function for extracting the raw data. Finally, the medical-grade ActiGraph wGT3X-BT (ActiGraph), often viewed as the gold standard in accelerometry movement analysis for research, features a built-in triaxial accelerometer that captures high-resolution raw acceleration data. The ActiLife software (ActiGraph LLC, version 6.11.4) was employed to initialize, process, and download data, extracting step counts and body posture measures (time spent lying down, sitting, and standing). In experiment 2, we used the ActiGraph wGT3X-BT as it was the selected device from experiment 1.

Protocols

Experiment 1

Patients from the first experiment were engaged in the activity data procedures to test the feasibility of 4 different wearable devices (ActiGraph wGT3X-BT, MOX1, MMC, and Fitbit Versa), as well as the concurrent validity of the ActiGraph in detecting body posture and step counts. Each patient wore all 4 devices on the waist, 3 devices on the thigh, and 3 on the ankle simultaneously and interchangeably, that is, up to 10 devices per patient. An elastic band equipped with Velcro on both ends was used to attach the devices securely to the body. The waistband featured 4 pockets, while the thigh and ankle bands had 3 pockets each, providing the flexibility to interchange the devices as needed during the study. The waistband was positioned at the level of the anterior superior iliac spine, the thigh band above the kneecap, and the ankle band above the malleolus. After positioning the elastic bands, the wearable devices were randomly assigned to each pocket. Once wearable devices were placed, patients performed the mobility protocol that included body posture tasks (standing, sitting, and lying) and the timed up and go (TUG) [18] mobility test.

During the body posture tasks, patients were asked to lie down, sit on the edge of the bed, and then stand with or without support for 5 minutes. The time spent in each body posture was observed and recorded by a physiotherapist. Following, the patients performed the TUG which is a reliable and valid test to assess mobility and balance in older adults [18]. It measures, in seconds, the time taken by an individual to stand up from a standard armchair, walk 3 meters, turn, walk back to the chair, and sit down. The participants wore their customary walking aid (none, cane, or walker), and no physical assistance was given. The recording initiated as the patient raised from the chair and stopped as the patient sat on the chair again. The time and the number of steps taken to complete the test were also counted and recorded. Following the activity procedures, patients were invited to complete the acceptability questionnaire (Textbox 1). The protocol took approximately 45 - 60 minutes to complete.

Textbox 1. Feasibility questionnaires for experiments 1 and 2.

Experiment 1

1. Have you ever used a device to measure physical activity in the past? (Responses: yes, no)
2. Would you be willing to wear the device for a longer period, 5 to 7 days, as part of a research study? (Responses: very likely, somewhat likely, not likely)
3. What part of the body would you prefer to wear the device? (Responses: waist, thigh, ankle)
4. Which of these devices would you likely use? (Responses: ActiGraph, MOX1, Versa, MetaMotionC)
5. How easy would it be for you to remember to use the device every day? (Responses: very easy, easy, very difficult, difficult)
6. Do you think this device would interfere with your daily routine? (Responses: no effect, minor effect, major effect)
7. Would you feel more motivated to move when wearing the device? (Responses: yes, no, no answer)

Experiment 2

1. Have you ever used a device to measure physical activity in the past? (Responses: yes, no)
   a. What devices did you use?
   b. What activity (or activities) did you track with the devices?
   c. How long ago was it that you used the devices?
   d. If you stopped using the devices, why did you stop?
2. Would you be willing to wear a device to measure physical activity for 5 to 7 days while in the hospital? (Responses: likely, uncertain, unlikely, very likely, very unlikely)
3. Would you be willing to wear a device daily to measure physical activity once you return home from the hospital, for a period of up to say 3 months? (Responses: yes, no)
4. What part of the body would you most prefer to wear the device? (Responses: wrist, waist, thigh, ankle, other)
5. Would you feel more motivated to move when wearing a device to measure physical activity? (Responses: yes, no)
6. Prior to being in the hospital, how many days per week on average did you engage in 30 minutes or more of physical activity, which was enough to raise your breathing rate? (Responses: 1, 2, 3, 4, 5, 6, 7, none)

Experiment 2

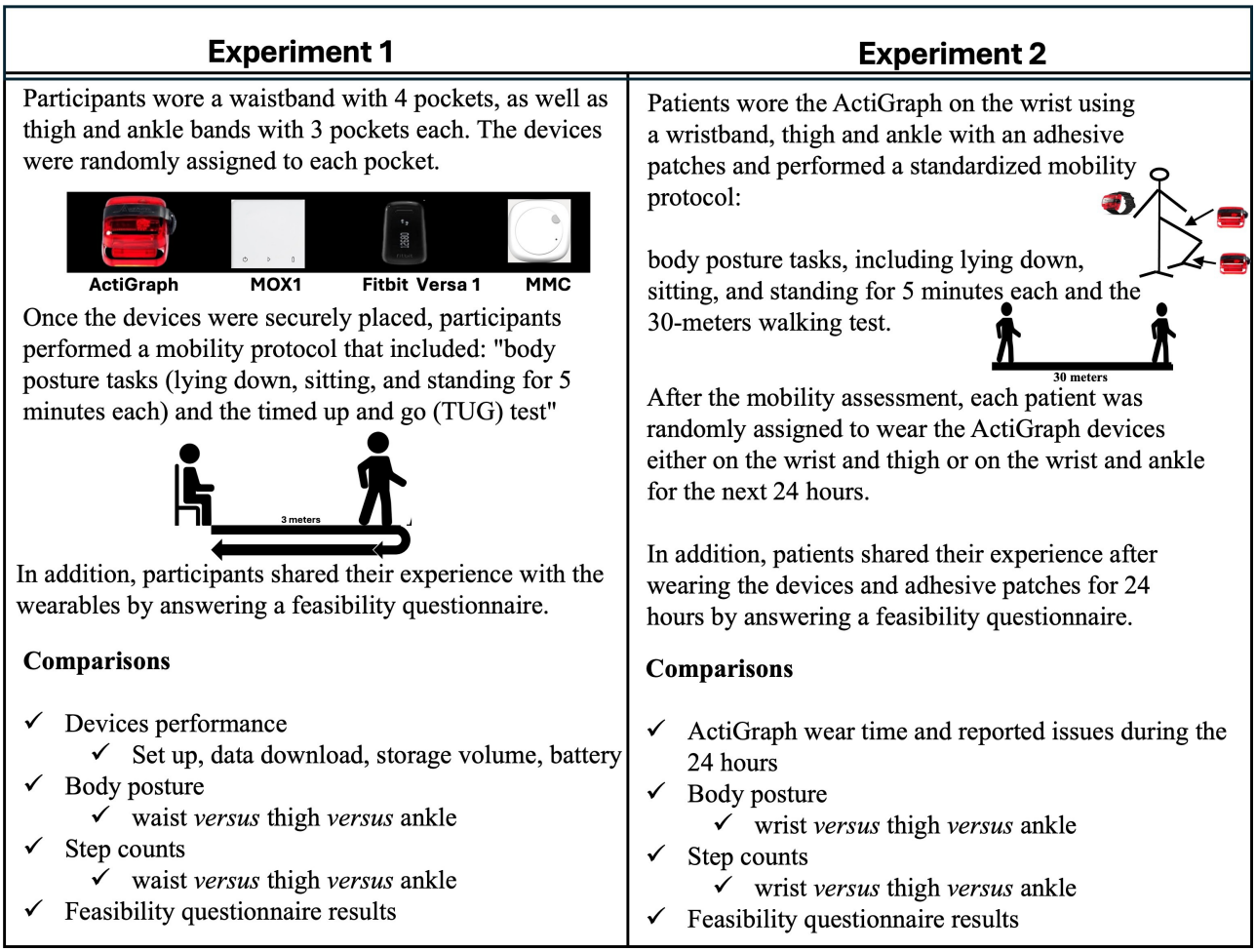
We conducted 1 mobility assessment session where patients wore the chosen device from experiment 1 (ActiGraph) in 3 different body locations (wrist, thigh, and ankle) and performed a standardized mobility protocol under the supervision of a trained physiotherapist. Patients wore the ActiGraph on the wrist using a wristband. At the thigh, the ActiGraph was attached to the anterior aspect of either the left or right thigh just above the kneecap, and at the ankle, the device was attached just above the malleoli. The Hypafix Stretch Non-Woven Adhesive (BSN Medical) was used to affix the devices. The

mobility protocol included the same body posture protocol as the first experiment, followed by a step count task recorded during a 30-meter walking test (30MWT) [19,20]. Patients were asked to stand up from a chair, walk 30 meter, turn around, and walk back at a comfortable pace using their usual walking aid if necessary. The time and the number of steps taken to complete the test was counted and recorded. The mobility protocol in experiment 2 lasted approximately 30 - 45 minutes.

After the mobility assessment, each patient was randomly assigned to wear the ActiGraph devices on either the wrist and thigh or wrist and ankle for the next 24 hours. We assessed the acceptability and feasibility of the wearable devices by location, as well as their acceptability of

wearing the adhesive patches over a 24-hour duration in the hospital. We also documented any interruptions or issues encountered during the 24-hour monitoring period. If patients removed or stopped using the device, the research assistant recorded feedback from both the patient and the attending nurses regarding the reasons for discontinuation. Figure 1 illustrated the variations within each experiment.

Figure 1. Summary of procedures for experiment 1 and 2.



Data Reduction

Experiment 1

Data were collected at 100 Hz for MOX1, and at 50 Hz for the other devices. In the second experiment, the ActiGraph data were collected at 30 Hz. Custom Python algorithms were created for the MMC and Fitbit devices to enable continuous saving and downloading of raw accelerometer data, necessitating the use of a companion device, a tablet. The MOX1 device was initialized using its proprietary software and data were downloaded using the provided MATLAB function. Raw data from these devices were not subjected to further processing; our focus was solely on recording downloading time, as well as the quantity and quality of the saved data.

In both experiments, the ActiLife software (version 6.11.4) was used to initialize, download, and process the ActiGraph data. Data were aggregated into 60 seconds time-stamped epochs. The following measures were obtained from the ActiLife: wear time, counts per minutes, activity intensity, step counts, and body posture (time spent in lying down, sitting, and standing). The following measures were obtained from the ActiLife algorithms: wear time, counts per minute, activity intensity, step counts, and body posture (time spent lying down, sitting, and standing). For step counts, since the ActiLife algorithms measures strides when attached to the thigh, we doubled the stride numbers to determine the actual number of steps taken. This adjustment is necessary because each stride recorded by the device corresponds to two steps—one for each leg—so the raw stride count is multiplied by 2 to obtain an accurate step count.

In addition, we applied the ActiGraph manufacturer's step algorithm, the low-frequency extension filter, to the step count data. This filter is designed to detect lower-amplitude movements, enhancing the accuracy of step detection [21]. Body posture classification in both experiments was obtained using the thigh-worn algorithm from the ActiGraph that relies on movement and the thigh angle to accurately classify lying and sitting versus standing positions [22]. The first and last 45 seconds of data from each activity were discarded to avoid potential participant error in recording time and to exclude the transition times.

Experiment 2

For the 24-hour protocol, the accelerometer data from the ActiGraph was screened for wear time using the method described by Choi et al [23]. Based on the activity counts determined by the ActiLife algorithms, and using an epoch length of 60 seconds, nonwear time was defined as 90 consecutive minutes of zero counts, with an allowance of 2 minutes of nonzero counts, provided there were 30-minute consecutive zero counts before and after that allowance. Based on the wear time, we determined the average amount of time the patients wore the devices at the wrist, thigh, and ankle during the 24-hour protocol.

Statistical Analysis

Descriptive data were analyzed using measures of central tendency and dispersion. Absolute percentage errors were calculated between the observed body posture time and step count against the values obtained using the ActiGraph algorithm (absolute percentage errors=[(observed data–ActiGraph data)÷observed data]×100). Intraclass correlation coefficients ($ICC_{2,1}$) [24] were used to examine criterion validity between step count taken from the ActiGraph compared with those observed during the TUG and 30MWT. By convention, an ICC ≥0.75 was considered excellent, 0.60 - 0.74 good, 0.40 - 0.59 fair, and <0.40 poor [25].

Bland-Altman plots [26,27] were constructed to show the variability of the ActiGraph in recording step count compared with the observed data. With this technique, the mean error score and the 95% prediction intervals can be examined in graphical form. Comparisons that are in closer agreement will have a mean bias close to zero and tighter 95% prediction intervals. Statistical analyses were conducted with SPSS version 26 (IBM Corp) with the α level set at .05.

Results

Overview

A total of 25 older adults (n=17, 65% women) with a mean age of 79.6 (SD 8.1) years participated in experiment 1, and 30 participants (n=24, 80% women), with a mean age of 81.4 (SD 8.8) years, participated in experiment 2 (Table 1). The patients were hospitalized due to the following conditions: cardiac conditions (n=6) including hypertension, congestive heart failure, atrial fibrillation, and coronary artery disease; urinary tract infection (n=5); deconditioning (n=5); falls (n=3); acute kidney injury (n=3); exacerbation of COPD (n=2); shortness of breath (n=2); sepsis (n=2); pneumonia (n=1); syncope (n=1); leg pain (n=1); vertigo (n=1); gastroenteritis (n=1); and back pain (n=1). Table 1 describes the patients' demographics and clinical characteristics. In general, patients were more independently mobile before admission.

Table . Demographics, comorbidities, and mobility characteristics of the patients.

Variables			Experiment 1	Experiment 2
Demographics, n			25	30
	Age (y), mean (SD)		79.6 (8.1)	81.4 (8.8)
	Sex (female), n (%)		17 (65)	24 (80)
	BMI (kg/m^2), mean (SD)		25.6 (6.2)	25.2 (7.7)
Gait aids, n (%)			25 (100)	29 (100)
	None		12 (48)	7 (24)
	Walker rollator		12 (48)	16 (55)
	Cane		1 (4)	2 (6)
	Cane and walker		0	4 (14)
Most common comorbidities, n			30	—[c]
	Cardiac condition (hypertension, congestive heart failure, atrial fibrillation, and coronary artery disease)		25	—
	Diabetes		9	—
	Osteoarthritis		8	—
	Cancer		7	—
	COPD[b] and Asthma		7	—
	Osteoporosis		5	—
	Kidney disease		5	—
	Early diagnosis of Alzheimer		2	—
	Parkinson		2	—
Mobility before admission, n (%)			—	30 (100)
	Independent		—	8 (27)
	Independent with cane		—	3 (10)
	Independent with walker		—	15 (50)
	1 Person assist		—	1 (3)
	1 Person assist with walker		—	2 (7)
	Independent with cane inside and walker outside		—	1 (3)
Mobility after admission			—	N=30
	Independent		—	1 (3)
	Independent with cane		—	1 (3)
	Independent with walker		—	11 (37)
	1 Person assist		—	3 (10)
	1 Person assist with walker		—	14 (47)
Mobility tests, n			23	—
	TUG[a] (s)			
		Mean (SD)	28.5 (13.7)	—
		Median (IQR)	23.3 (7.14-61.0)	—
	Average speed (m/s)			
		Mean (SD)	0.27 (0.16)	—
		Median (IQR)	0.22 (0.09-0.84)	—
	30MWT[d] (m-s) test 1			

Variables		Experiment 1	Experiment 2
	Mean (SD) (s)	—	50 (31)
	Median (IQR) (s)	—	35 (30-120)
	Percentile 25 (s)	—	32
	Percentile 75 (s)	—	67
30MWT (m-s) retest			
	Mean (SD) (s)	—	58 (17)
	Median (IQR) (s)	—	20 (30-50)
	Percentile 25 (s)	—	50
	Percentile 75 (s)	—	70

[a]TUG: timed up and go.

[b]COPD: chronic obstructive pulmonary disease.

[c]Not applicable.

[d]30MWT: 30-meter walking test.

Experiment 1

The acceptability questionnaire showed that out of 25 patients, 24 patients (96%) reported no previous experience with wearables. However, 20 patients (80%) expressed their willingness to wear a wearable for a period of 5-7 days. In terms of the preferred body location, the ankle was selected by 14 patients (58%), the waist was selected by 7 patients (29%), and the thigh by 4 patients (17%). Furthermore, 1 patient did not answer this question. Among the wearable devices, the Fitbit Versa was selected by 12 patients (48%), the ActiGraph was chosen by 6 patients (24%), and the MOX1 and MMC were selected by 3 (12%) and 2 (8%) patients, respectively.

We observed that both the ActiGraph and MOX1 devices were easy to set up and enabled faster data downloads compared with Fitbit and MMC devices. Of the 4 devices, only the ActiGraph retrieved 100% of the collected data. The MATLAB function from the MOX1 retrieved 72% of the files, with 1 file containing missing data, and 7 files having a lower frequency than the specified 100 Hz setup. In addition, the MATLAB function provided by the manufacturer did not include time stamps. The algorithms on the Fitbit Versa were able to retrieve 79% of the data, with the primary issue being missing data. Similarly, the MMC algorithms retrieved 81% of the data, although some data were missing due to a time stamp error that was not identified until later in the data collection process. In addition to its high performance, the ActiGraph device also had the longest battery life and storage volume compared to all other devices. The proprietary software offered by ActiGraph allowed us to process, quickly view and extract the data using a comprehensive selection of independently developed and validated algorithms. Given the ease of data collection using the ActiGraph in the above experiment, it was selected as the most user-friendly device for further evaluation.

For the body posture assessments, the waist- and thigh-worn ActiGraph identified the lying down position correctly 73.6% and 78.2% of the time, respectively. For the standing posture, both the thigh- and ankle-worn ActiGraph achieved high identification rates of 83.8% and 82.3%, respectively. However, all devices exhibited poor performance in identifying the sitting position, ranging from 25.4% to 49.6% (Table 2).

Table . Percentage of times the ActiGraph correctly detected body posture compared with the physiotherapist recordings during experiment 1 (n=25).

Patient body posture	Experiment 1 devices' attachment								
	Waist, %			Thigh, %			Ankle, %		
	Lie[a]	Sit	Stand	Lie	Sit	Stand	Lie	Sit	Stand
Lie	74[b]	0	0	78[b]	54	0	54[b]	0	0
Sit	11	41[b]	38	0	25[b]	15	0	50[b]	15
Stand	0	57	61[b]	1	1	84[b]	0.1	48	84[b]
Off[c]	16	1	1	21	20	1	45	3	3

[a]Lie: laying down.

[b]Percentage of time the ActiGraph device correctly identified the body postures.

[c]Off: percentage of time that the device detected it was off.

Table 3 displays the results of the ICC analysis, comparing the step counts during the TUG tests recorded by the ActiGraph devices worn on the waist, thigh, and ankle compared with direct observation by the physiotherapist. Data from 2 participants were excluded due to errors in registering the start and end of the tests. The ankle-worn device demonstrated the highest agreement with the physiotherapist ($ICC_{2,1}$=0.94, 95% CI 0.85-0.97), the lowest bias (average of the mean difference=0.9 steps), and a lower percentage of error counting steps (12.8%). The waist-worn device also shows excellent agreement with direct observation ($ICC_{2,1}$=0.85, 95% CI 0.65-0.94) but higher bias (1.4 steps) and a higher percentage of error counting steps (21%). The thigh-worn device has the lowest agreement ($ICC_{2,1}$=0.75, 95% CI −0.21 to 0.93), the highest bias (overcounted on average 7.3 steps), and the highest percentage of error counting steps (28.8%).

Table . Intraclass correlations and percentage of error in counting steps between the physiotherapist (observer) step count and ActiGraph during the timed up and go test in experiment 1 (n=23).

Body location	Bias	ICC[a]	95% CI	*P* value	Percentage error counting steps
Waist	1.4	0.85	0.65 to 0.94	.001	21
Thigh	7.3	0.75	−0.21 to 0.93	.001	28.8
Ankle	0.9	0.94	0.85 to 0.97	.001	12.8

[a]ICC: intraclass correlation coefficient.

Experiment 2

The acceptability questionnaire was completed by 25 patients. Their responses indicated that 22 (88%) had never used a wearable before, 19 (76%) would wear a device for 5 to 7 days while in hospital, and 16 (64%) were willing to wear the device daily at home, up to 3 months, as part of a research study. Furthermore, 12 (48%) patients felt motivated to move when wearing a device, and 8 (32%) would prefer to wear the device on the wrist. Intercurrences during the 24-hour protocol, include the following challenges with the devices: 1 participant had the thigh device removed and reapplied by nurses, while another participant removed both the ankle and wrist devices due to discomfort and itching. In 1 case, the wrist device was taken off because of itchiness, and another participant forgot the purpose of the devices, leading to the removal of both the wrist and thigh devices. In addition, the wrist device was loosened because of swelling, 1 participant had to remove the ankle device for 24 hours due to a major infection on the lower leg, and another was unable to wear the thigh device because the tape caused irritation.

Of the 30 patients, the ActiGraph devices were positioned on the wrist and ankle in 11 (37%) patients and on the wrist and thigh in 19 (63%) patients for the 24-hour protocol. One thigh-worn device recorded only 1:36 of data during the 24-hour protocol so the data were not included in the analysis. On average, the patients wore the device on the wrist for 22:45 (range 20:58-25:51), on the thigh for 24:36 (range 23:33-26:51), and on the ankle for 20:11 (range 8:03-26:16). Intercurrences reported while wearing the devices included itchiness at the wrist and thigh, patients removing devices due to forgetting their purpose, and patients loosening the device due to joint swelling.

The ActiGraph algorithms for the thigh-worn devices have recently changed, combining lying down and sitting as sedentary posture and including stepping detection (Table 4). The thigh-worn devices identified 100% of sedentary posture while the patient was lying down, 98% while they were sitting, and 91% while they were standing. The ankle-worn devices best identified lying (89%) and standing (84%) postures, and poorly identified the sitting posture (43.2%). In addition, the ankle-worn devices classified the position as sedentary on average 93% of the time when lying, but only 49% while the patients were seated. The wrist-worn devices performed poorly compared with the thigh- and ankle-worn devices, identifying lying down between 49% and 52% of the time and sitting and standing around 15% - 25% of the time. The wrist-worn devices were able to identify 80% of the sedentary posture while the patient was lying down, and 71% while sitting (Table 4).

Table . Percentage of times the ActiGraph correctly detected body posture compared with the physiotherapist recordings during experiment 2 (N=30).

Patient's body posture	Experiment 2 devices' attachment										
	Wrist, %			Thigh, %				Ankle, %			
	Sed[a]	Lie[b]	Sit	Stand	Sed	Stand	Step	Sed	Lie	Sit	Stand
Lie	84	50[c]	34	8	100[c]	0	0	92	89[c]	3	1.3
Sit	72	46	26[c]	23	98[c]	1	1	47	4	43[c]	49
Stand	76	10	66	24[c]	8	91[c]	0	13	0	13	84[c]
Off[d]	—	8	5	0	—	—	—	—	6	4	2

[a]Sed: sedentary (lying down + sitting).

[b]Lie: laying down.

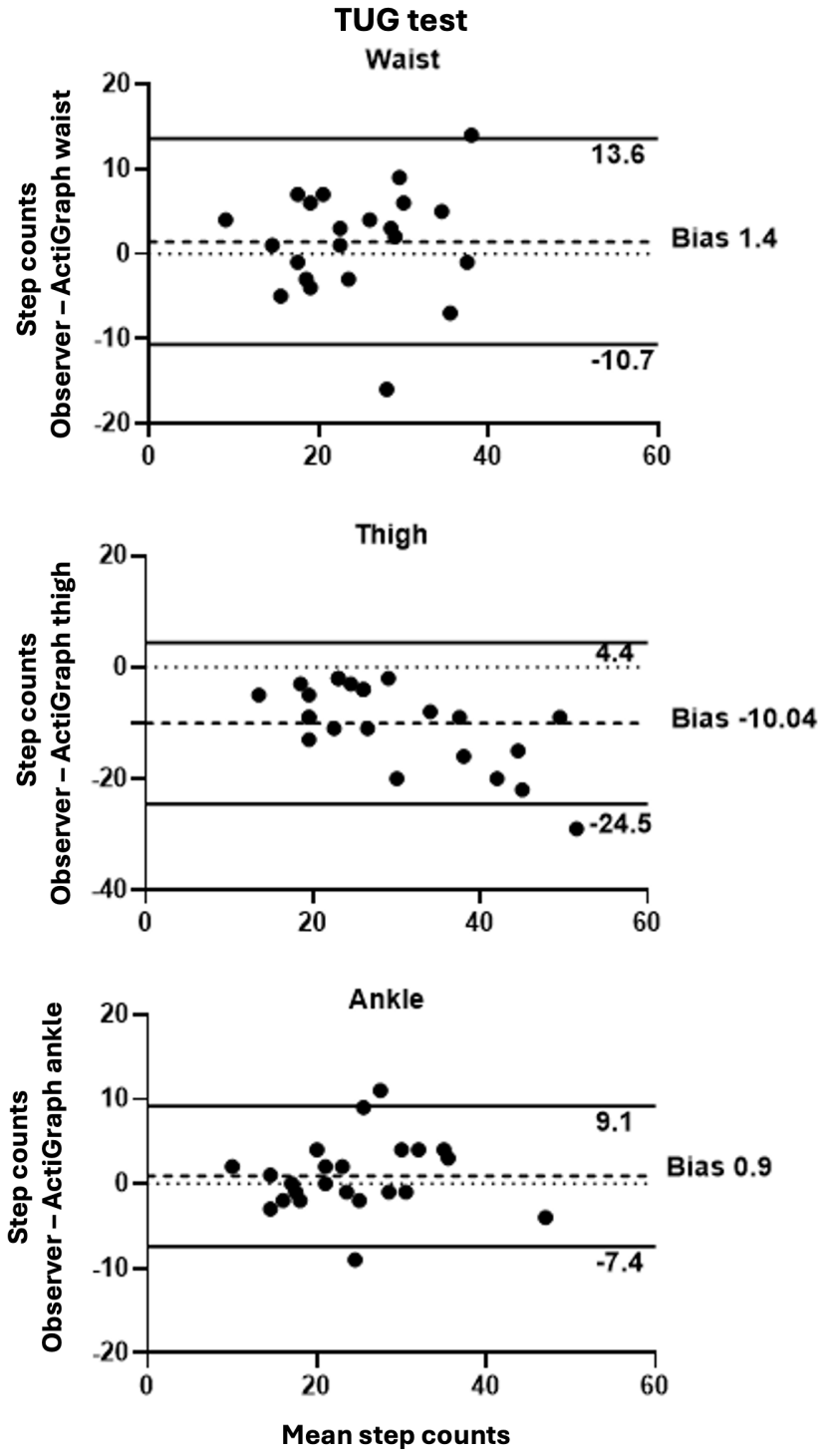
[c]Percentage of time the ActiGraph device correctly identified the body postures.

[d]Off: percentage time that the device detected it was off.

[e]Not applicable

Figure 2 shows the Bland-Altman plots of the difference between the observed step count and the ActiGraph placed on the waist, thigh, and ankle during the TUG test during experiment 1. Visual interpretation of the plots shows less bias (0.9), narrower limits of agreement (−7.4 to 9.1), and more observations closer to zero for the ankle-worn ActiGraph. In contrast, the thigh-worn device showed a larger limit of agreement, suggesting poor agreement between the 2 measurements, and a tendency to overcount steps.

Figure 2. Bland-Altman plots comparing observer measurements with ActiGraph data from devices worn on the waist, thigh, and ankle during the TUG test in experiment 1. The dashed lines denote the 95% limits of agreement (SD 1.96) of the mean difference, with darker dashed lines highlighting the mean difference or bias. TUG: timed up and go.



During the 30MWT in experiment 2, the wrist- and thigh-worn devices showed poor ICC values and a high percentage of miscounted steps. The ankle-worn devices showed excellent reliability and on average overcounted the number of steps by 1.9% (Table 5).

Table . Intraclass correlation coefficients and absolute percentage error of steps between the physiotherapist step count and the ActiGraph during the 30-Minute Walk Test in experiment 2 (n=30).

Body location	Bias	ICC[a]	95% CI	*P* value	Percentage error counting steps
Wrist	–42.1	–0.032	–0.217 to 0.218	.62	49.4
Thigh	–36.0	0.331	–0.149 to 0.706	<.001	56.8
Ankle	1.6	0.959	0.915 to 0.981	<.001	1.9

[a]ICC: intraclass correlation coefficient.

The Bland-Altman plots in Figure 3 illustrate the agreement between observer-rated step counts and ActiGraph step counts at the wrist, thigh, and ankle locations during the 30MWT. Among the 3 locations, the ankle-worn device demonstrated the best agreement with observer ratings, exhibiting a small bias and most observations falling within the 95% limits of agreement. In contrast, the wrist-worn device exhibited poor agreement, with the largest bias and widest limits of agreement.

Figure 3. Bland-Altman plots comparing observer measurements with ActiGraph data from devices worn on the wrist, thigh, and ankle during the 30 MWT in experiment 2. The dashed lines denote the 95% limits of agreement (SD 1.96) of the mean difference), with darker dashed lines highlighting the mean difference or bias. 30 MWT: 30-meter walk test.

Discussion

This study aimed to identify the optimal wearable device and wear location to assess and monitor mobility among older patients during hospitalization. We observed a high level of acceptability and feasibility regarding the usability and accuracy of wearable devices for detecting and monitoring activity in older patients hospitalized for acute medical illness. Although most patients had limited experience with activity monitoring devices, they were willing to wear one while hospitalized. Patients showed a preference for the Fitbit and ActiGraph devices, worn at the ankle or waist. Both the ActiGraph and Mox1 devices were easy to set up, but only the ActiGraph allowed for the retrieval of all data collected. Overall, the ActiGraph wGT3X-BT emerged as the preferred device with superior usability, data acquisition, and management.

The ActiGraph wGT3X-BT is widely recognized as the gold standard research-grade device for wearable mobility tracking, which is consistent with the findings of this study in our population. Its previous use in hospital settings underscores its utility for monitoring physical activity and posture in clinical contexts [28-30]. For example, other studies have employed the ActiGraph to predict hospital-acquired disability [28] and demonstrated its accuracy in quantifying postures and activity levels among hospitalized adults [29]. This study supports the device's reliability and validity among older hospitalized patients with unique challenges such as slower movement patterns and increased sedentary behavior.

However, when planning a cohort study in a hospital setting, it is crucial to consider multiple factors beyond just device accuracy, including accessibility to raw data and patient acceptability. During our study, the Fitbit Versa 1 was a popular device choice but lacked the option to download raw data directly. Instead, the download process relied on a companion device, a tablet, and any miscommunication between the 2 devices resulted in data lost. The MMC was selected for its affordability and open-source platform. However, we encountered challenges with the downloading process, which was time-consuming, and we also noted timestamp errors. The MATLAB function provided by the MOX 1 also resulted in missing data and lack of time stamps. Despite the challenges encountered with the MMC, MOX 1, and Fitbit devices in our study, it is plausible that these issues have since been addressed and resolved, underscoring the potential advancements made by these companies in their software and device functionalities. Our study highlights the importance of considering broader factors, such as data accessibility and device functionality, when evaluating the feasibility of wearable technology for hospital-based cohort studies.

Studies have suggested the lack of physical activity and immobilization during hospitalization may be more related to aspects of hospital care rather than to the patient's diagnosis [1,31]. While performance-based mobility tests can predict functional decline and hospital discharge, they are not commonly integrated into hospital measures [32-34]. Conducting a gait speed test, for example, is time-consuming and might not be feasible on a day-to-day basis in the hospital setting. Therefore, the use of wearable technology could be an attractive option, requiring minimal time investment for both patients and health professionals. There is a substantial body of literature on the utilization of wearables in hospitalized patients [16,28,34,35]. However, the research on patient feedback regarding their experiences and perceptions of wearing the devices while hospitalized is limited [17], which is crucial for optimizing their use in health care settings. In addition, we were particularly concerned about the potential interference of these devices with patients' medical conditions. Prolonged sitting and fluid intake, common in hospitalized patients, can lead to swelling, particularly in the ankles and wrists. We observed instances of wrist and ankle swelling in patients that resulted in the devices being removed, underscoring the importance of considering these issues when implementing wearable technology in hospital settings. Furthermore, we also considered the possibility of the devices interfering with medical equipment commonly used during hospitalization, including colostomy bags, urinary catheters, and wound dressings, as any disruption to these essential devices could compromise patient care and safety. For instance, in our first experiment, we applied an elastic band with pockets to the patient's waist, thigh, and ankle. We noticed that the waistband was problematic because it would interfere with heart sensor wires and colostomy bags. In addition, the thigh elastic band would easily fall from the participant's leg. Thus, in the second experiment, we disregarded the waistband and used an adhesive patch, which could also bring discomfort due to skin itchiness, as observed in our study. Considering these factors before launching a larger cohort is crucial and might save time and effort for both patients and researchers.

Our findings highlight the importance of considering both practical and contextual factors when selecting wear locations for mobility monitoring in hospital settings. While the literature supports the thigh for measuring postural behaviors and the waist or ankle for step counts, these locations may not always be feasible in clinical environments. Wrist-worn devices, although convenient and widely accepted, presented challenges such as patient discomfort due to swelling and skin irritation. This study serves as a preliminary exploration of acceptable and practical wear locations in a hospital setting, emphasizing the need to balance feasibility with the specific mobility metrics of interest.

It is well-established that sedentary behavior in hospitalized older adults is associated with a heightened risk of hospital-acquired disability (HAD) and functional decline [28,29]. Therefore, accurately measuring sedentary behavior in this population is essential for timely intervention and management strategies within the hospital setting. Body posture poses a challenge in accurate measurement, ideally requiring the use of at least 2 devices. For instance, thigh-worn devices have been reported as optimal for placing accelerometers to determine sedentary behavior (ie, lying and sitting). In this regard, the thigh-worn inclinometer algorithm provided by ActiGraph uses a movement threshold and thigh angular orientation to distinguish lying and sitting from standing and stepping [36]. Compared with other algorithms such as the activPAL, the ActiGraph's thigh angular parameter improves the classification of sitting posture, even when the participant

has their legs crossed or stretched, in both laboratory and free-living conditions [22,37]. Waist-worn devices are effective for differentiating lying from sitting but not from sitting and standing, while ankle-worn devices distinguish lying and standing. In our study, we tested the ActiGraph on the waist, thigh, and ankle, and in experiment 1, all wear locations performed well for lying and standing but poorly for sitting. We believe that the thigh elastic band likely changed position when participants transitioned from lying to sitting, changing the orientation of the ActiGraph, thus affecting the algorithms' ability to detect sitting posture. In experiment 2, we incorporated a wrist-worn device due to its user-friendly nature. To mitigate the issue encountered in experiment 1 with the elastic band, we adopted adhesive patches to secure the devices on the thigh and ankle. The thigh device exhibited the highest performance across various postures, followed by the ankle device, which showed adequate accuracy in detecting lying and standing postures. However, the wrist device, despite its ease of use, performed poorly in detecting all body postures. Our findings are supported by the literature [17,29], and we recommend using an accelerometer on the thigh and ankle to capture more detailed information on sedentary behavior.

Our investigation on step counts showed that the ankle is the most accurate body position to capture this metric. This finding aligns with previous research examining the validity of ankle-worn ActiGraph devices in conjunction with the lower frequency extension filter for step counting among hospitalized older adults [16,38]. In their study, Webber and St John [38] demonstrated that the ActiGraph positioned on the ankle was comparable with direct observation (ICC=0.94, median absolute error=2.5%) for monitoring step counts during the 10-meter walk test in hospitalized older adults. In addition, Anderson et al [16] reported that the ActiGraph positioned on the ankle (mean difference=−0.85 steps, ICC=0.99) accurately records step counts in hospitalized adults during free self-selected walking. By leveraging the ankle and thigh as the placement site for capturing mobility, researchers and health care practitioners can enhance the accuracy of monitoring and promote more effective interventions aimed at improving mobility and overall health outcomes in hospitalized adults and similar populations.

A potential limitation of this study is the bias introduced by patients wearing multiple devices simultaneously in experiment 1. The discomfort or inconvenience of wearing several devices may have influenced their feedback, as they were likely focused on the overall experience rather than evaluating each device individually. However, the primary goal of experiment 1 was to assess the feasibility of using multiple devices in a hospital setting, with a focus on understanding the practical and real-world challenges of device wearability, data retrieval, and integration in a clinical environment. While this may introduce some bias in the acceptability results, it does not diminish the value of these insights into how devices function together in practice.

Building on these findings, our study has several strengths, including identifying optimal devices and placement sites for wearable data collection, as well as validation through comparison with a gold-standard observer. However, it is important to note that our free-living protocol in experiment 2 was limited to a 24-hour duration. Despite this constraint, we were able to gather valuable insights, particularly regarding the feasibility and integration of wearable technology in hospital settings.

In conclusion, our study found that the ActiGraph wGT3X-BT was the most feasible device for assessing and monitoring mobility among older hospitalized patients. The ActiGraph's thigh-worn algorithm accurately detects sedentary behavior under supervised conditions and, when paired with a device at the ankle, provides detailed information on lying and sitting postures. In addition, our findings indicate that step counts can be accurately detected using the low-frequency extension with devices on the ankle. Therefore, we recommend the use of two devices, at the thigh and ankle, to accurately measure sedentary behavior and step count among older people in a hospital setting. Longer-term studies are warranted to evaluate the use of wearable data for predicting health outcomes after hospitalization and for informing clinical decision-making and efforts to promote early mobility among older hospitalized patients.

Acknowledgments

This project was supported by grants from the McMaster Institute for Research on Aging, McMaster University. MB holds a Tier2 Canada Research Chair in Mobility, Aging and Chronic Disease (CRC-2020-00043) and an Early Researcher Award from the Government of Ontario. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

LG, MD, and MB contributed to the conception and design of the work. RK, CC, and PN contributed to the acquisition and analysis of data. PN and RK contributed to the interpretation of data. YH, RZ, and SR contributed to the development of the software for data analysis of the MetaMotionC wearable. All authors contributed to drafting the work (or substantial revisions). All authors have read and approved the final manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Krumholz HM. Post-hospital syndrome--an acquired, transient condition of generalized risk. N Engl J Med 2013 Jan 10;368(2):100-102. [doi: 10.1056/NEJMp1212324] [Medline: 23301730]
2. Brown CJ, Foley KT, Lowman JD Jr, et al. Comparison of posthospitalization function and community mobility in hospital mobility program and usual care patients: a randomized clinical trial. JAMA Intern Med 2016 Jul 1;176(7):921-927. [doi: 10.1001/jamainternmed.2016.1870] [Medline: 27243899]
3. Creditor MC. Hazards of hospitalization of the elderly. Ann Intern Med 1993 Feb 1;118(3):219-223. [doi: 10.7326/0003-4819-118-3-199302010-00011] [Medline: 8417639]
4. Brown CJ, Redden DT, Flood KL, Allman RM. The underrecognized epidemic of low mobility during hospitalization of older adults. J Am Geriatr Soc 2009 Sep;57(9):1660-1665. [doi: 10.1111/j.1532-5415.2009.02393.x] [Medline: 19682121]
5. Jawad BN, Petersen J, Andersen O, Pedersen MM. Variations in physical activity and sedentary behavior during and after hospitalization in acutely admitted older medical patients: a longitudinal study. BMC Geriatr 2022 Mar 15;22(1):209. [doi: 10.1186/s12877-022-02917-8] [Medline: 35291952]
6. Rojer AGM, Ramsey KA, Trappenburg MC, et al. Patterns of objectively measured sedentary behavior and physical activity and their association with changes in physical and functional performance in geriatric rehabilitation inpatients. J Am Med Dir Assoc 2023 May;24(5):629-637. [doi: 10.1016/j.jamda.2023.01.011] [Medline: 36841261]
7. Callen BL, Mahoney JE, Grieves CB, Wells TJ, Enloe M. Frequency of hallway ambulation by hospitalized older adults on medical units of an academic hospital. Geriatr Nurs (Lond) 2004 Jul;25(4):212-217. [doi: 10.1016/j.gerinurse.2004.06.016]
8. Reynolds CD, Brazier KV, Burgess EAA, et al. Effects of unstructured mobility programs in older hospitalized general medicine patients: a systematic review and meta-analysis. J Am Med Dir Assoc 2021 Oct;22(10):2063-2073. [doi: 10.1016/j.jamda.2020.12.008] [Medline: 33434569]
9. Trost SG, O'Neil M. Clinical use of objective measures of physical activity. Br J Sports Med 2014 Feb;48(3):178-181. [doi: 10.1136/bjsports-2013-093173] [Medline: 24311601]
10. Ngueleu AM, Barthod C, Best KL, Routhier F, Otis M, Batcho CS. Criterion validity of ActiGraph monitoring devices for step counting and distance measurement in adults and older adults: a systematic review. J Neuroeng Rehabil 2022 Oct 17;19(1):112. [doi: 10.1186/s12984-022-01085-5] [Medline: 36253787]
11. Berendsen BAJ, Hendriks MRC, Meijer K, Plasqui G, Schaper NC, Savelberg HHCM. Which activity monitor to use? Validity, reproducibility and user friendliness of three activity monitors. BMC Public Health 2014 Jul 24;14:749. [doi: 10.1186/1471-2458-14-749] [Medline: 25059233]
12. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159. [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
13. Van Remoortel H, Giavedoni S, Raste Y, et al. Validity of activity monitors in health and chronic disease: a systematic review. Int J Behav Nutr Phys Act 2012 Jul 9;9(1):84. [doi: 10.1186/1479-5868-9-84] [Medline: 22776399]
14. Karaca A, Demirci N, Yılmaz V, Hazır Aytar S, Can S, Ünver E. Validation of the ActiGraph WGT3X-BT accelerometer for step counts at five different body locations in laboratory settings. Meas Phys Educ Exerc Sci 2022 Jan 2;26(1):63-72. [doi: 10.1080/1091367X.2021.1948414]
15. Baldwin C, van Kessel G, Phillips A, Johnston K. Accelerometry shows inpatients with acute medical or surgical conditions spend little time upright and are highly sedentary: systematic review. Phys Ther 2017 Nov 1;97(11):1044-1065. [doi: 10.1093/ptj/pzx076] [Medline: 29077906]
16. Anderson JL, Yoward LS, Green AJ. A study investigating the validity of an accelerometer in quantification of step count in adult hospital inpatients recovering from critical illness. Clin Rehabil 2019 May;33(5):936-942. [doi: 10.1177/0269215519829893] [Medline: 30764647]
17. Hartley P, Keevil VL, Westgate K, et al. Using accelerometers to measure physical activity in older patients admitted to hospital. Curr Gerontol Geriatr Res 2018;2018:3280240. [doi: 10.1155/2018/3280240] [Medline: 30420882]
18. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
19. Andersson M, Moberg L, Svantesson U, Sundbom A, Johansson H, Emtner M. Measuring walking speed in COPD: test-retest reliability of the 30-metre walk test and comparison with the 6-minute walk test. Prim Care Respir J 2011 Dec;20(4):434-440. [doi: 10.4104/pcrj.2011.00082] [Medline: 21938352]
20. Bohm PE, Fehlings MG, Kopjar B, et al. Psychometric properties of the 30-m walking test in patients with degenerative cervical myelopathy: results from two prospective multicenter cohort studies. Spine J 2017 Feb;17(2):211-217. [doi: 10.1016/j.spinee.2016.08.033] [Medline: 27592193]
21. Feito Y, Hornbuckle LM, Reid LA, Crouter SE. Effect of ActiGraph's low frequency extension for estimating steps and physical activity intensity. PLoS One 2017;12(11):e0188242. [doi: 10.1371/journal.pone.0188242] [Medline: 29155845]
22. Steeves JA, Bowles HR, McClain JJ, et al. Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. Med Sci Sports Exerc 2015 May;47(5):952-959. [doi: 10.1249/MSS.0000000000000497] [Medline: 25202847]
23. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364. [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]

24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
25. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. Am J Ment Defic 1981 Sep;86(2):127-137. [Medline: 7315877]
26. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015;25(2):141-151. [doi: 10.11613/BM.2015.015] [Medline: 26110027]
27. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb;327(8476):307-310. [doi: 10.1016/S0140-6736(86)90837-8] [Medline: 2868172]
28. Pavon JM, Sloane RJ, Pieper CF, et al. Accelerometer-measured hospital physical activity and hospital-acquired disability in older adults. J Am Geriatr Soc 2020 Feb;68(2):261-265. [doi: 10.1111/jgs.16231] [Medline: 31747050]
29. Anderson JL, Green AJ, Yoward LS, Hall HK. Validity and reliability of accelerometry in identification of lying, sitting, standing or purposeful activity in adult hospital inpatients recovering from acute or critical illness: a systematic review. Clin Rehabil 2018 Feb;32(2):233-242. [doi: 10.1177/0269215517724850] [Medline: 28805075]
30. Kirk AG, Kimmel LA, Behm KJ, Peiris CL, Ekegren CL. Validity of the activPAL and ActiGraph for measuring sitting time and steps in hospitalised orthopaedic patients with altered weight bearing. Disabil Rehabil 2024 Jan;46(2):378-386. [doi: 10.1080/09638288.2022.2157896] [Medline: 36541196]
31. Zisberg A, Syn-Hershko A. Factors related to the mobility of hospitalized older adults: a prospective cohort study. Geriatr Nurs 2016;37(2):96-100. [doi: 10.1016/j.gerinurse.2015.10.012] [Medline: 26597674]
32. Ibrahim K, Owen C, Patel HP, et al. Can routine clinical data identify older patients at risk of poor healthcare outcomes on admission to hospital? BMC Res Notes 2017 Aug 10;10(1):384. [doi: 10.1186/s13104-017-2705-7] [Medline: 28797300]
33. Kumar A, Karmarkar AM, Graham JE, et al. Comorbidity indices versus function as potential predictors of 30-day readmission in older patients following postacute rehabilitation. J Gerontol A Biol Sci Med Sci 2017 Feb;72(2):223-228. [doi: 10.1093/gerona/glw148] [Medline: 27492451]
34. Fisher SR, Graham JE, Ottenbacher KJ, Deer R, Ostir GV. Inpatient walking activity to predict readmission in older adults. Arch Phys Med Rehabil 2016 Sep;97(9 Suppl):S226-S231. [doi: 10.1016/j.apmr.2015.09.029] [Medline: 27264549]
35. van Roekel EH, Winkler EAH, Bours MJL, et al. Associations of sedentary time and patterns of sedentary time accumulation with health-related quality of life in colorectal cancer survivors. Prev Med Rep 2016 Dec;4:262-269. [doi: 10.1016/j.pmedr.2016.06.022] [Medline: 27419042]
36. How is inclination determined (for thigh wear location)? ActiGraph. URL: https://actigraphcorp.my.site.com/support/s/article/How-is-Inclination-Determined-for-Thigh-Wear-Location [accessed 2023-03-12]
37. Edwardson CL, Rowlands AV, Bunnewell S, et al. Accuracy of posture allocation algorithms for thigh- and waist-worn accelerometers. Med Sci Sports Exerc 2016 Jun;48(6):1085-1090. [doi: 10.1249/MSS.0000000000000865] [Medline: 26741122]
38. Webber SC, St John PD. Comparison of ActiGraph GT3X+ and StepWatch step count accuracy in geriatric rehabilitation patients. J Aging Phys Act 2016 Jul;24(3):451-458. [doi: 10.1123/japa.2015-0234] [Medline: 26751505]

Abbreviations

30MWT: 30-meter walking test
ICC: intraclass correlation coefficient
MMC: MetaMotionC
TUG: timed up and go

Edited by J Wang; submitted 21.07.24; peer-reviewed by A Habib, E Cenko; revised version received 13.12.24; accepted 30.01.25; published 12.05.25.

Please cite as:
Nascimento P, Kirkwood R, Griffith LE, Duong M, Cooper C, Hao Y, Zheng R, Raza S, Beauchamp M
Identifying Optimal Wearable Devices for Monitoring Mobility in Hospitalized Older Adults: Feasibility, Acceptability, and Validity Study
JMIR Aging 2025;8:e64372
URL: https://aging.jmir.org/2025/1/e64372
doi:10.2196/64372

Performance of a Digital Cognitive Assessment in Predicting Dementia Stages Delineated by the Dementia Severity Rating Scale: Retrospective Study

Duong Huynh[1], PhD; Kevin Sun[1], BS; Mary Patterson[1], BA, MS; Reza Hosseini Ghomi[1,2,3], MD, MSE; Bin Huang[1], PhD

[1]BrainCheck, Inc, Austin, TX, United States

[2]Frontier Psychiatry, PLLC, Billings, MT, United States

[3]Department of Neurology, Institute for Neuroengineering, & eScience Institute, University of Washington, Seattle, WA, United States

Corresponding Author:
Bin Huang, PhD
BrainCheck, Inc, Austin, TX, United States

Abstract

Background: Dementia is characterized by impairments in an individual's cognitive and functional abilities. Digital cognitive assessments have been shown to be effective in detecting mild cognitive impairment and dementia, but whether they can stage the disease remains to be studied.

Objective: In this study, we examined (1) the correlation between scores obtained from BrainCheck standard battery of cognitive assessments (BC-Assess), a digital cognitive assessment, and scores obtained from the Dementia Severity Rating Scale (DSRS), and (2) the accuracy of using the BC-Assess score to predict dementia stage delineated by the DSRS score. We also explored whether BC-Assess can be combined with information from the Katz Index of Independence in activities of daily living (ADL) to obtain enhanced accuracy.

Methods: Retrospective analysis was performed on a BrainCheck dataset containing 1751 patients with dementia with different cognitive and functional assessments completed for cognitive care planning, including the DSRS, the ADL, and the BC-Assess. The patients were staged according to their DSRS total score (DSRS-TS): 982 mild (DSRS-TS 10 - 18), 656 moderate (DSRS-TS 19-26), and 113 severe (DSRS-TS 37-54) patients. Pearson correlation was used to assess the associations between BC-Assess overall score (BC-OS), ADL total score (ADL-TS), and DSRS-TS. Logistic regression was used to evaluate the possibility of using patients' BC-OS and ADL-TS to predict their stage.

Results: We found moderate Pearson correlations between DSRS-TS and BC-OS (r=−0.53), between DSRS-TS and ADL-TS (r=−0.55), and a weak correlation between BC-OS and ADL-TS (r=0.37). Both BC-OS and ADL-TS significantly decreased with increasing severity. BC-OS demonstrated to be a good predictor of dementia stages, with an area under the receiver operating characteristic curve (ROC-AUC) of classification using logistic regression ranging from .733 to .917. When BC-Assess was combined with ADL, higher prediction accuracies were achieved, with an ROC-AUC ranging from 0.786 to 0.961.

Conclusions: Our results suggest that BC-Assess could serve as an effective alternative tool to DSRS for grading dementia severity, particularly in cases where DSRS, or other global assessments, may be challenging to obtain due to logistical and time constraints.

(JMIR Aging 2025;8:e65292) doi:10.2196/65292

KEYWORDS

stage; severity; progression; correlation; association; cognitive impairment; functional activities; cognitive assessment; BrainCheck; dementia; Alzheimer disease; gerontology; geriatric; old; elderly; aging; retrospective analysis; digital assessment; patient assessment; digital cognitive assessment; digital health; neurodegeneration; memory loss; memory function; risk factors

Introduction

Dementia is characterized by impairments in an individual's cognitive and everyday functional abilities. Although the pattern and advancement of these impairments vary among patients, the disease is usually considered as having 3 main stages: mild (early-stage), moderate (middle-stage), and severe (late-stage) [1]. The distinction between these categories lies in the extent to which a patient's physical, cognitive, and psychosocial well-being deteriorates. The rate of deterioration widely varies and progresses through subtle changes in daily functioning to complete loss of independence and the need for a caregiver [1]. Staging of dementia has important implications for practical

decision-making and research [2]. From the practical standpoint, knowledge of disease severity is helpful for selection of appropriate intervention options, for prognosis and communication with patients and their family about expectations, care needs, as well as early planning for the future [3-5]. From the research standpoint, staging patients is needed to determine their eligibility for participation, to achieve better clinical efficacy, and to obtain homogeneous sampling in research studies [4], particularly clinical trials.

As noted earlier, dementia is not a monolithic disease; therefore, both cognitive and functional abilities need to be measured to accurately assess its severity and progression. Standardized cognitive tests can provide objective measures of cognitive functioning in different domains such as memory and executive function. Brief cognitive screening tests such as the Montreal Cognitive Assessment and the Mini-Mental State Examination are widely used in clinical practices, but they are rarely used to grade dementia severity [6,7]. Formal neuropsychological tests provide a more comprehensive evaluation of cognitive functioning to support differential diagnoses [8,9]. However, these tests typically take hours and require administration by specialists. Digital cognitive assessments are emerging as an efficient solution due to their self-administration capability, remote accessibility, and automated scoring. Although these types of assessments have been shown to be effective in detecting mild cognitive impairment and dementia [10-12], their ability to stage dementia is not yet clear. Functional assessments quantify the ability to perform activities of daily living through questionnaires such as the Katz Index of Independence in activities of daily living (ADL) [13,14]. Functional assessments are valuable for helping to evaluate dementia severity and also for providing proper guidance to patients and their caregivers.

In clinical practice, both cognitive and functional deficits can be measured using global staging scales. These scales typically come in the form of a questionnaire or interview, relying on subjective judgments and reports from patients or their knowledgeable informants. Compared with other scales, such as the Global Deterioration Scale [15] or the Functional Assessment Staging [16], the Clinical Dementia Rating (CDR) [17-19] scale appears to be the most well-studied and best-evidenced for dementia staging [20,21]. However, the use of this instrument is often impractical in many situations where time and cost are concerns, due to its semi-structured nature, long administration time (30 - 60 min), and requirement of clinical judgment from a trained professional during administration and scoring [18,22]. In response to this, brief instruments have been developed to mirror the CDR, including the Dementia Severity Rating Scale (DSRS) [23]. The DSRS uses a multiple-choice format in which the caregiver rates the patient's cognitive and functional ability in 12 categories [23,24]. The DSRS has been shown to be effective in staging and determining the progression of dementia [23,25,26].

While cognitive tests have been shown to correlate with the above dementia staging tools [4,27,28], previous research primarily focused on traditional paper-based cognitive tests. The increasing adoption of digital solutions and tools in health care calls for the re-evaluation of dementia staging tools and digital cognitive assessments. The first goal of this study was to examine the correlation between scores obtained from the BrainCheck standard battery of cognitive assessments (BC-Assess), a digital cognitive assessment, and scores obtained from the DSRS, a global staging scale. The second goal of this study is to evaluate the effectiveness of using the BC-Assess score to predict dementia stage delineated by the DSRS score. We also explored whether BC-Assess can be combined with information from ADL to obtain enhanced predictive capability.

Methods

Data Source

This retrospective study analyzed a real-world dataset of patients who received cognitive care planning services from their providers through BrainCheck Plan. These patients and their caregivers had completed a series of assessments, including DSRS, ADL, and BC-Assess, and received a comprehensive and personalized cognitive care plan. Inclusion criteria for this study were (1) patients 60 years of age or older; (2) assessments of cognitive care planning completed in English on an iPad; and (3) Dementia Severity Rating Scale total score (DSRS-TS) ≥10. The criterion for DSRS-TS was to only include patients that were rated by DSRS to have mild, moderate, and severe dementia [29].

Patients that showed evidence of moderate to severe depression, defined by a Geriatric Depression Scale score of 9 or above [30], were excluded. This is to avoid including reversible causes of dementia, which may have poorly correlated impacts on cognitive and functional measures. Given that depression is common among patients with dementia [31-33], patients with mild depression were not excluded to avoid overfiltering of the data. For patients having multiple care plans, only the latest record was considered. To reduce the impact of practice effect, only data from providers who had completed ≥20 cognitive care planning services for their patients were included. In total, data from 1751 patients with their cognitive care plans completed between February 2022 and April 2024 across 48 providers were included for this analysis.

Measurements

The DSRS is a brief informant-based questionnaire made up of 12 items that measure functional abilities, including memory, orientation to time, orientation to place, decision-making, social interaction, home activities, personal care, eating, toileting, mobility, recognition, and speech and language. DSRS-TS is calculated by adding up scores across 12 functional areas, ranging from zero (no impairment) to 54 (extreme impairment) [25]. The patients could be categorized into 3 groups of different dementia severity levels based on their DSRS-TS. The majority 56% (n=982) were mild-stage patients (DSRS-TS 10 - 18), moderate-stage patients (DSRS-TS 19 - 26) accounted for 37.5% (n=656), and the remaining 6.5% (n=113) were severe-stage patients (DSRS-TS 37 - 54). The 3 severity levels serve as class labels in a logistic regression analysis in this study to predict patients' conditions based on their BC-Assess and ADL data.

The BC-Assess, completed by the patients, consisted of 6 individual cognitive assessments: Immediate Recognition and

Delayed Recognition (memory), Digit Symbol Substitution (processing speed), Stroop (executive function), Trails Making Test A, and Trails Making Test B (attention or mental flexibility). Detailed descriptions of these tests can be found in previous studies [12,34]. The raw score for each assessment is calculated using assessment-specific measurements based on accuracy or reaction time (Table S1 in Multimedia Appendix 1). The BC-Assess raw overall score is the average of raw scores from all completed assessments after each score has been transformed from its natural range into a common range [0,100] using the formula in Table S1 in Multimedia Appendix 1. A *z*-score is then calculated for each assessment score and for the overall score using the corresponding age- and device-specific mean and standard deviation values from the BrainCheck normative database. Assessment standardized scores and BC-Assess overall standardized score (BC-OS) are reported by rescaling the *z*-scores to have a mean of 100 and a standard deviation of 15.

The ADL is an informant-based 6-item survey that measures an individual's ability to independently perform basic activities of daily living, including bathing, dressing, going to the toilet, transferring, continence, and feeding. It is calculated by adding up scores from the 6 categories, each of which takes a value of 1 for independent and 0 for dependent, resulting in an ADL total score (ADL-TS) ranging from 0 to 6. An ADL-TS of 2 or less indicates severe functional impairment, 3 - 4 indicates moderate impairment, and 5 - 6 indicates full function [13,35].

Statistical Analysis

Data analyses were performed using Python (version 3.8.5). Descriptives were presented for demographics and each score. The χ^2 test was used to examine whether the distribution of gender was similar in patients from the 3 groups. The Kruskal-Wallis test was used to compare the mean age of patients across groups.

A 3-way multivariate analysis of variance (MANOVA) was used to examine the joint variation of ADL-TS and BC-OS as a function of dementia stage, age group, and gender. Post-hoc analysis employed 1-way MANOVAs to compare these 2 scores across dementia stages for each individual age and gender group. Tests for statistical significance of the mean differences across stages, age groups, and genders were also performed separately for each score using 3-way ANOVAs. For these statistical comparisons, age is treated as a factor with 3 levels: 60 - 69, 70 - 79, and 80+.

Logistic regression was used to investigate the effectiveness of using the patients' BC-OS and ADL-TS to predict their dementia stage, where age and gender served as covariates. In this analysis, age is treated as a continuous variable, and gender is treated as a binary variable: 1 for male and 0 for female. Although the 3 dementia stages form ordinal classes, separate binary logistic regressions were used to classify mild versus moderate; moderate versus severe; and mild versus severe, because the proportional odds assumption was not satisfied for both BC-OS (P=.01) and ADL-TS (P<.001) from Brant's Wald test [36], suggesting ordered logistic regression was not appropriate.

The binary logistic regression model for predicting a patient's condition is:

$$(1)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta2\cdot\text{stADL}+\beta3\cdot\text{stAge}+\beta4\cdot\text{stGender}$$

where p is the probability of predicting the patient as having a pre-specified positive class. In each of the above binary classifications, we chose the more severe stage to be the positive class. stBC,stADL,stAge,stGender are standardized values of predictor variables BC-OS, ADL-TS, age, and gender. The coefficients βi`s (i=1 - 4) represent the effects of the 4 predictors, and β0 is the intercept.

Model fitting was based on weighted loss functions to take into consideration class imbalance across the 3 dementia groups. Model performance was evaluated using 5-fold cross-validation with stratification, repeated 20 times (100 iterations), such that on each iteration, all training and testing subsets had roughly the same proportion of patients from each group as in the original dataset. We compared four different models:

(1) full model that included all 4 predictors as in equation (1)

(2) reduced model 1 that included BC-OS and ADL-TS:

$$(2)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta2\cdot\text{stADL}$$

(3) reduced model 2 that included BC-OS, age, and gender:

$$(3)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}+\beta3\cdot\text{stAge}+\beta4\cdot\text{stGender}$$

(4) reduced model 3 that included only BC-OS:

$$(4)\text{logit}(p)=\beta0+\beta1\cdot\text{stBC}$$

An receiver operating characteristic (ROC) curve was generated for each model on each cross-validation iteration. Paired-samples *t*-tests were used to compare areas under the receiver operating characteristic curves (ROC-AUC) across models.

Ethical Considerations

This study was conducted using existing deidentified data collected through BrainCheck. The dataset contained no personal identifiers, and no attempt was made to reidentify the individuals. As such, the research does not meet the definition of "human subjects research" as outlined by HHS 45 CFR 46.102. Therefore, this study did not require ethics review or approval by an institutional review board.

Results

Demographics and Assessment Performance

Table 1 summarizes the demographic characteristics of the patients in this study. Group sample size decreased with increasing severity for both genders. The distribution of gender was similar across the 3 groups (P=.84). Although the range of age was similar, mean age significantly increased with severity (P<.001; pairwise comparisons: P<.001 for mild vs moderate and mild vs severe, P=.02 for moderate vs severe). Statistical comparisons were not performed for education level and race due to a lot of missing information.

Table . Demographics and summary statistics of scores across dementia severity groups and the total sample.

Demographic characteristics	Mild group (n=982)	Moderate group (n=656)	Severe group (n=113)	Total (N=1751)
Gender, n (%)[a]				
Female	589 (60)	394 (60.1)	71 (62.8)	1054 (60.2)
Male	393 (40)	262 (39.9)	42 (37.2)	697 (39.8)
Age, years[b]				
Mean (SD)	78 (8.3)	80.3 (8.4)	80.9 (9.0)	79.0 (8.6)
Range	60 - 101	60 - 102	61 - 101	60 - 102
Age bucket (years), n (%)[c]				
60 - 69	161 (16.4)	74 (11.3)	15 (13.3)	250 (14.3)
70 - 79	393 (40)	232 (35.4)	34 (30.1)	659 (37.6)
≥80	428 (43.6)	350 (53.3)	64 (56.6)	842 (48.1)
Education (years), n (%)				
>12	72 (7.3)	49 (7.5)	1 (0.9)	122 (7)
≤12	48 (4.9)	41 (6.3)	9 (8)	98 (5.6)
Not reported	862 (87.8)	566 (86.3)	103 (91.2)	1531 (87.4)
Race, n (%)				
White	108 (11)	72 (11)	6 (5.3)	186 (10.6)
Black	23 (2.4)	12 (1.8)	2 (1.8)	37 (2.1)
Others	19 (1.9)	17 (2.6)	1 (0.8)	37 (2.1)
Not reported	832 (84.7)	555 (84.6)	104 (92.0)	1491 (85.2)
DSRS-TS[d], mean (SD)	13.4 (2.5)	25.4 (4.8)	43.3 (5.0)	19.8 (9.2)
BC-OS[e], mean (SD)	62.2 (35.1)	30.1 (38.9)	–5.0 (26.8)	45.8 (41.4)
ADL-TS[f], mean (SD)	4.8 (1.7)	3.2 (2.1)	1.2 (1.5)	4.0 (2.1)

[a]P=.84 (χ^2 test): distribution of gender was not significantly different across groups.

[b]P<.001 (Kruskal-Wallis test): mean age of patients was significantly different across groups.

[c]P<.001 (χ^2 test): distribution of age bucket was significantly different across groups.

[d]DSRS-TS: Dementia Severity Rating Scale total score.

[e]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[f]ADL-TS: activities of daily living total score.

The means and standard deviations of DSRS-TS, ADL-TS, and BC-OS are provided in Table 1. Overall, the BC-OS and ADL-TS decreased with increasing severity delineated by the DSRS-TS. Figure 1 shows the distributions of BC-OS and ADL-TS across patients within each group. For both scores, the distribution was systematically skewed toward the high values for the mild group, toward the low values for the severe group, and more evenly distributed for the moderate group. Table 2 shows the number and percentage of cases in each group that fall in different BC-Assess and ADL-TS score intervals.

Figure 1. Box plots of the distributions of BrainCheck standard battery of cognitive assessments overall score (BC-OS; left) and activities of daily living total score (ADL-TS; right) for each patient group: green=mild, yellow=moderate, and red=severe. Normally distributed random noise was used to add displacements along the x-axis for patients within each group (left and right) and along the y-axis at each ADL-TS value (right).

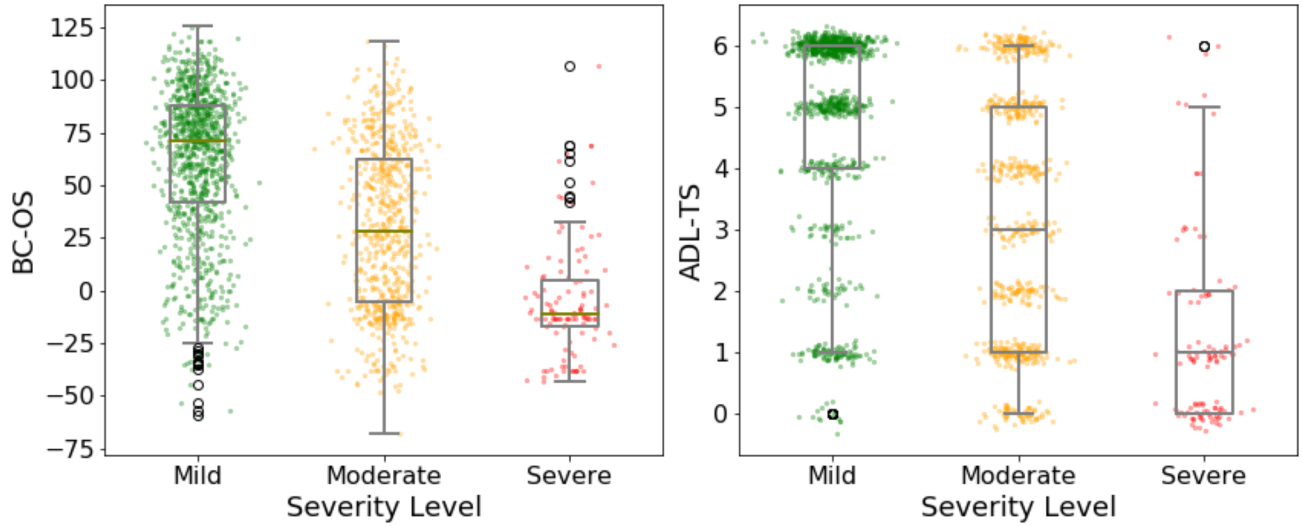


Table . ADL-TS[a] and BC-OS[b] distributions across dementia severity groups.

Score	Mild group (n=982)	Moderate group (n=656)	Severe group (n=113)
BC-OS, n (%)			
Beyond 2 SD of normative mean[c]	480 (48.9)	532 (81.1)	112 (99.1)
Within 2 SD of normative mean	502 (51.1)	124 (18.9)	1 (0.9)
ADL-TS, n (%)			
0 - 2 (severe)	137 (14)	275 (41.9)	96 (85)
3 - 4 (moderate)	119 (12.1)	142 (21.7)	10 (8.8)
5 - 6 (Full function)	726 (73.9)	239 (36.4)	7 (6.2)

[a]ADL-TS: activities of daily living total score.

[b]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

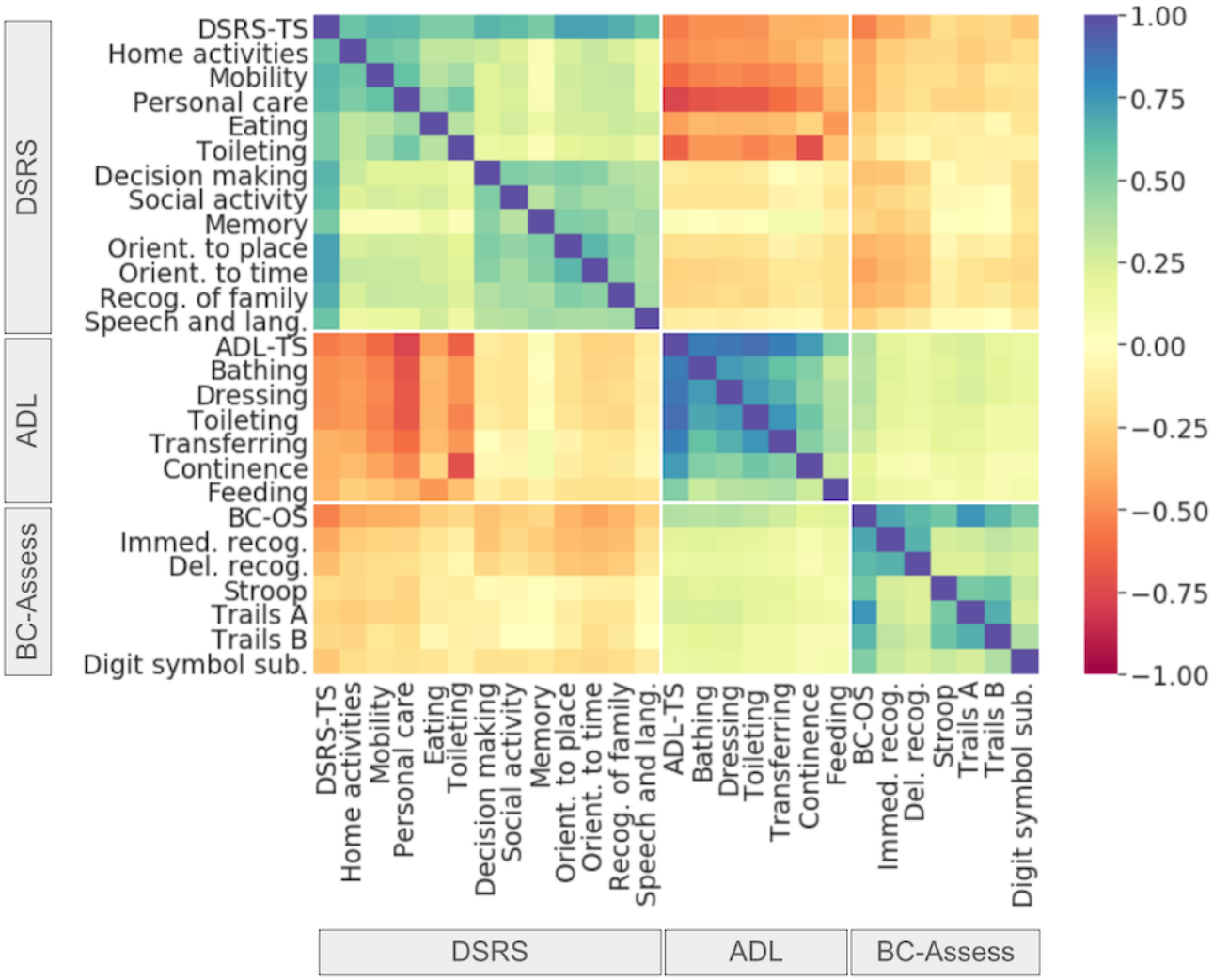
[c]Based on a BC-OS normative mean of 100 and a standard deviation of 15.

Correlations Between Assessments

We found moderate Pearson correlations between DSRS-TS and BC-OS (r=−0.53; P<.001), between DSRS-TS and ADL-TS (r=−0.55; P<.001), and a weak correlation between BC-OS and ADL-TS (r=0.37; P<.001). Since DSRS covers both cognitive and functional performance of a patient, moderate associations between DSRS-TS with BC-OS and ADL-TS were as expected. The weak correlation between ADL-TS and BC-OS suggests that cognitive and functional abilities are associated with each other, but certain discrepancies exist between the two.

The heatmap in Figure 2 plots Pearson correlations across DSRS, BC-Assess, and ADL subscores. Compared with BC-Assess subscores, BC-OS showed stronger correlations with DSRS and ADL subscores. However, these correlations were weak. Among DSRS subscores, BC-OS was most associated with home-activities (r=−0.41; P<.001), mobility (r=−0.39; P<.001), personal-care (r=−0.39; P<.001), orientation-to-time or orientation-to-place (r=−0.43 and −0.37; P<.001), and recognition-of-family (r=−0.37; P<.001). Among ADL subscores, BC-OS was most associated with dressing (r=0.36; P<.001), bathing (r=0.35; P<.001), and toileting (r=0.33; P<.001). Although weaker, a clear trend can be observed from Figure 2 for BC-Assess subscores. With regard to the DSRS, assessments of memory (Immediate or Delayed Recognition) were most associated with memory-demanding activities such as memory, orientation-to-time, orientation-to-place, recognition-of-family, and decision-making, whereas assessments of executive function, attention, or mental flexibility (Stroop, Trails Making A/B) were most associated with home-activities, mobility, and personal-care. With regard to the ADL, BC-Assess subscores were more associated with bathing, dressing, and toileting than with categories that are more essential, such as feeding, continence, and transferring. Between ADL and DSRS subscores, correlations mainly occurred within a subset of DSRS activities that are of the same types as those rated by the ADL such as eating, home-activities, mobility, personal-care, and toileting. Of these, the strongest correlations were found between DSRS toileting and ADL continence (r=−0.72; P<.001), and between DSRS personal-care and ADL bathing (r=−0.70; P<.001), dressing (r=−0.68; P<.001), and toileting (r=−0.68; P<.001).

Figure 2. Correlations between BC-Assess (BrainCheck standard battery of cognitive assessments), DSRS (Dementia Severity Rating Scale), and ADL (activities of daily living) subscores. DSRS-TS: DSRS total score; ADL-TS: ADL total score; BC-OS: BC-Assess overall score.



Comparison of BC-OS and ADL-TS Across Dementia Stages, Age Groups, and Genders

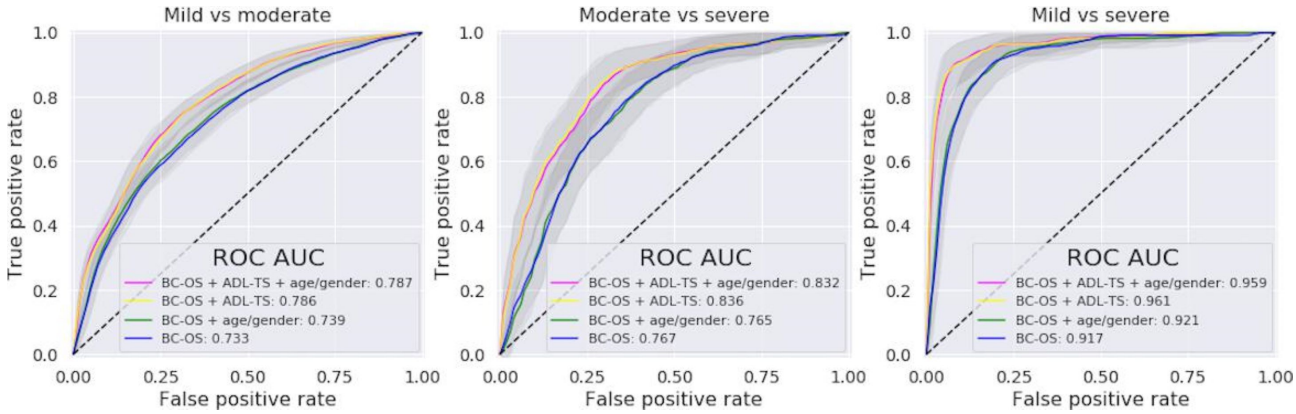
Given the correlation between BC-OS and ADL-TS, we analyzed the differences in these two scores across dementia stages, age groups, and genders by running a 3-way MANOVA (*BC-OS + ADL-TS~ Dementia Stage * Age Group * Gender*). Results based on the Pillais Trace method showed a significant effect of Dementia Stage (Pillais Trace =0.046 , $F_{43,466}$=20.5; *P*<.001) whereas the effects of age and gender and all interaction terms were not significant. Post hoc 1-way MANOVAs (*BC-OS + ADL-TS~ Dementia Stage*) were run to compare BC-OS and ADL-TS combined between mild versus moderate and between moderate versus severe separately for each age and gender group. Except for the 60 - 69 and female group (n=136) showing a nonsignificant difference between moderate versus severe, significant differences were observed for all other cases. We further performed 3-way ANOVAs where BC-OS (*BC-OS~ Dementia Stage * Age Group * Gender*) and ADL-TS (*ADL-TS~ Dementia Stage * Age Group * Gender*) were considered separately. For BC-OS, we found a significant main effect of Dementia Stage only ($F_{21,733}$=270.31; *P*<.001). The insignificant differences in BC-OS across age groups were as expected as this score had been standardized to adjust for age differences. For ADL-TS, we found significant effects of Dementia Stage ($F_{21,733}$=278.87; *P*<.001), Age Group ($F_{21,733}$=4.77; *P*=.009), and Gender ($F_{11,733}$=7.82; *P*=.005). For both scores, no interaction terms were significant.

Logistic Regression to Examine the Roles of BC-OS, ADL-TS, Age, and Gender in Predicting a Patient's Condition

ROC analysis (Figure 3) showed a comparable performance between the full model (*BC-OS + ADL-TS + age or gender*: ROC-AUC=0.787 for mild vs moderate; 0.832 for moderate vs severe; and 0.959 for mild vs severe) and reduced model 1 (*BC-OS + ADL-TS*: ROC-AUC=0.786 for mild vs moderate; 0.836 for moderate vs severe; and 0.961 for mild vs severe), and between reduced model 2 (*BC-OS + age or gender*: ROC-AUC=0.739 for mild vs moderate; 0.765 for moderate vs severe; and 0.921 for mild vs severe) and reduced model 3 (*BC-OS* only: ROC-AUC=0.733 for mild vs moderate; 0.767 for moderate vs severe; and 0.917 for mild vs severe). The small differences in ROC-AUC generated by the omission of age and gender suggest that they are not important predictors. Moreover, including these demographic factors appears to have led to some degree of overfitting where reduced model 1 performed slightly but significantly better than the full model (mild vs severe:

P=.002; moderate vs severe: *P*<.001) and reduced model 3 performed slightly but significantly better than reduced model 2 (mild vs moderate: *P*<.001; moderate vs severe: *P*=.04; mild vs severe: *P*<.001).

Figure 3. Model comparisons for the classification between mild and moderate (left) and between moderate and severe (right): receiver operating characteristic curves and mean area under the receiver operating characteristic curve (ROC AUC) for the full model (BrainCheck standard battery of cognitive assessments overall score [BC-OS] + activities of daily living total score [ADL-TS] + age and gender; magenta), reduced model 1 (BC-OS + ADL-TS; yellow), reduced model 2 (BC-OS + age and gender; green), and reduced model 3 (BC-OS only; blue). The shaded area along each curve represents the corresponding standard deviation across cross validation iterations. In each classification, the more severe condition was chosen to be the positive class.



The 2 models that include ADL-TS (full model and reduced model 1) performed significantly better than the 2 models without ADL-TS (reduced models 2 and 3) (*P*<.001 for each pairwise comparison), suggesting the important role of ADL. For all binary classifications, although BC-OS alone can serve as a fairly effective predictor with an ROC-AUC of at least 0.733, including ADL-TS in the model significantly improved prediction accuracy.

As age and gender were not important factors, we excluded them from further analysis. We examined the diagnostic performance of reduced models 1 (*BC-OS + ADL-TS*) and 3 (*BC-OS* only) at the optimal cut-points from their ROC curves. These are points on the ROC curves that maximize true positive rate and minimize false positive rate. Table 3 shows sensitivity (true positive rate) and specificity (1–false positive rate) at the optimal cut-point for each model and classification. When included, ADL-TS improved the sensitivity and specificity by 3% - 5%. For reduced model 3, which included only BC-OS, we found that the optimal cut-point corresponded to a BC-OS of 52 (roughly 3 standard deviations below the normative mean) for the classification between mild versus moderate, and a BC-OS of 0 for the classification between moderate versus severe.

Table . Sensitivity (Se) and specificity (Sp) by model and classification: mean (SD) across cross-validation iterations. In each classification, the more severe condition was chosen to be the positive class.

Classification	BC-OS[a] + ADL-TS[b]		BC-OS only	
	Se, mean (SD)	Sp, mean (SD)	Se, mean (SD)	Sp, mean (SD)
Mild versus moderate	0.74 (0.04)	0.72 (0.04)	0.69 (0.05)	0.68 (0.05)
Moderate versus severe	0.84 (0.06)	0.76 (0.05)	0.77 (0.07)	0.72 (0.07)
Mild versus severe	0.92 (0.04)	0.93 (0.03)	0.89 (0.04)	0.86 (0.04)

[a]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[b]ADL-TS: activities of daily living total score.

The fitted model coefficients are provided for reduced models 1 and 3 in Table 4. The 1-sided *P* value obtained from bootstrapping for each coefficient is shown in the parentheses.

Table . Coefficients (*P* values) of the fitted reduced models 1 (*BC-OS*[a] + *ADL-TS*[b]) and model 3 (*BC-OS* only).

	Regression coefficients, β (*P* value)				
	Reduced model 1			Reduced model 3	
Classification	β0 (Intercept)	β1 (stBC)	β2(stADL)	β0 (Intercept)	β1 (stBC)
Mild versus moderate	−0.071 (.01)	−0.757 (<.001)	−0.717 (<.001)	−0.053 (.002)	−0.890 (<.001)
Moderate versus severe	−0.819 (<.001)	−0.989 (.03)	−1.059 (<.001)	−0.453 (<.001)	−1.215 (<.001)
Mild versus severe	−2.443 (<.001)	−1.611 (<.001)	−1.517 (<.001)	−1.557 (<.001)	−2.076 (<.001)

[a]BC-OS: BrainCheck standard battery of cognitive assessments overall score.

[b]ADL-TS: activities of daily living total score.

Discussion

By conducting a retrospective analysis of patient data in real-world clinical settings, this study sought to investigate the relationship between cognitive performance in a battery of digital cognitive assessments, BC-Assess, and dementia severity measured by the DSRS. We found a statistically significant moderate correlation between the BC-OS and the DSRS-TS. This correlation is comparable with that between the ADL-TS and the DSRS-TS. Both BC-OS and ADL-TS significantly decrease with increasing severity. BC-Assess demonstrated to be a good predictor of dementia severity, with ROC-AUC of classification using logistic regression ranging from 0.733 to 0.917. When BC-Assess was combined with ADL, higher prediction accuracies were achieved, with ROC-AUC ranging from 0.786 to 0.961.

Our results suggest that BC-Assess could serve as an alternative tool to DSRS for grading dementia severity, particularly in cases where DSRS, or other global assessments, may be challenging to obtain due to logistical and time constraints. Unlike DSRS, BC-Assess, as a brief digital cognitive assessment, offers the advantage of flexible choice of self-administration or administration by clinical support staff, runs on common consumer technology, and does not require availability of an informant. The significant improvement of prediction accuracies when BC-Assess is combined with ADL indicates the synergetic relationship between cognitive and functional measures in grading dementia severity. Previous studies have shown that patients' loss of independence to manage activities of daily living is nonlinearly related with their cognitive decline [37], and their correlation weakens as the disease progresses [38]. This is consistent with the finding of relatively low correlation (r=0.37) between the 2 measures in this study and elsewhere [39,40]. Together, these findings suggest that ADL carries additional information of functional abilities that is partially independent of cognitive abilities measured by BC-Assess. When combined, the 2 measures provide a more comprehensive understanding of a patient's condition.

While the BC-OS and the ADL-TS from the mild and severe groups separate well from each other, scores from the moderate group vary widely among patients and largely overlap with both the mild and severe groups. This is reflected in high sensitivity and specificity (.86 or higher) for the classification between the mild and severe groups, and moderate sensitivity and specificity (0.83 or lower) for the classifications of the moderate group. Overall, however, a gradual change in the distribution of each score across stages can be observed. In line with current knowledge about the progression and stages of dementia [1,41], this pattern of results suggests that cognitive and functional declines do not happen in the same way across patients with dementia, and that there might only be subtle differences in cognitive or functional performance, or both, across patients within 2 successive stages.

Implicit in this study is the assumption that the staging of dementia severity based on the DSRS-TS had accurately identified each patient's underlying degree of severity. Previous studies demonstrated that the DSRS has high test-retest and inter-rater reliability [25] and good concurrent validity with high correlations with the CDR, the Mini-Mental Status Examination, and the Consortium to Establish a Registry for Alzheimer's Disease [23,25]. Other studies showed that the DSRS-TS can effectively differentiate between individuals with dementia, MCI, and healthy controls [42], and that it changes at a constant linear rate throughout the entire clinical course of dementia [25]. However, the psychometric properties of the DSRS in distinguishing between patients with mild, moderate, and severe dementia have not been studied. The DSRS-TS cut-offs used for staging of dementia severity have been suggested based on the mapping of the DSRS-TS onto CDR stages where a CDR global score of 0, 0.5, 1, 2, and 3 represent no, questionable, mild, moderate, and severe dementia. However, it has been shown that there is a large variability in the degree of severity among patients placed in a particular CDR stage, and patients with the same degree of severity can be placed in different CDR stages [2,43,44]. Furthermore, the precision of severity grading depends on the scoring approach to the CDR, ie, the item-response-theory approach is more precise than the sum-of-the-boxes approach, which is more precise than the global score approach [43]. On top of that, the mapping of DSRS-TS to CDR global score was based on a limited sample of patients with dementia that might not be representative of patients at different severity stages in general [23]. These limitations are possible contributing factors to the widespread distributions and overlaps of BC-OS and ADL-TS across patient groups delineated by the DSRS-TS in this study. To allow for more systematic investigations into the effectiveness of using these scores in dementia staging, future research needs to address these limitations and establish more fine-grained and well-defined staging criteria as well as optimize

the scoring methods for the DSRS, the CDR, and other assessments of dementia severity.

Suboptimal and inconsistent data quality in a dataset acquired outside typical clinical research settings is another factor that potentially causes large variabilities of scores observed in this study [45]. While real-world data may better represent diverse clinical environments, which is desired to obtain a generalized relationship between assessments, limited control over the data collection process and differences in clinical practices may result in reduced accuracy and consistency of the data. Inconsistencies exist across clinical sites and across units within each site due to differences in internal policies, staff training, workflow, and expertise. Inconsistencies also come from the different patient or caregiver populations each site serves. Patients may also have different dementia pathologies, leading to significant differences in the pattern of scores.

The ADL and BC-Assess also have their own limitations that one should take into consideration when interpreting the current results. The ADL measures 6 basic activities of daily living and employs dichotomous scoring, which allows only 2 possible scores for each functional category, ie, 1 for *independent* and 0 for *dependent*. Therefore, it lacks the resolution to capture intermediate levels of dependence. Furthermore, as it is subjective ratings from informants, for cases with small changes in functional activities, patients may end up receiving substantially different scores depending on their caregivers' judgments, resulting in low interrater variability [46,47]. In this study, with only patients with dementia included (based on their DSRS-TS), we found high variability in the ADL-TS across patients within each group, especially for the moderate group. As for the BC-Assess, besides measurement errors that exist in any assessment and technical difficulties older adults may have when using smart devices for the assessment, it might have limited utility in severity staging because patients with extremely severe conditions might not be able to complete it [48], and it could also suffer from a floor effect, a common limitation of psychometric tests [49].

Our data show a high imbalance in the number of patients across the 3 groups, with mild, moderate, and severe dementia accounting for 56%, 37.5%, and 6.5%, respectively. Although the trend is similar, higher proportions of patients estimated to be in the moderate (31%) and severe (21%) stages of Alzheimer's Disease were reported in a previous study [50]. As patients included in this study were those that received cognitive care planning services from their providers through the BrainCheck Plan platform within a 2-year period, our data do not necessarily reflect the prevalence of each stage. The fact that we only included patients with DSRS, ADL, and BC-Assess data also limited the number of patients in later stages who might be too impaired to take the BC-Assess. Furthermore, given its main goals are to help patients and their families better understand the patients' condition and needs, to offer strategies to improve their overall quality of life, and to plan for the future when their condition gets worse, cognitive care planning is more meaningful for patients in early stages. Patients in the severe stage of dementia are completely dependent on their family or caregivers, and many of them require specialized care and attention in facilities. These institutionalized patients tend to have been diagnosed and given care plans tailored to their specific needs by the institution.

In addition to the findings of this study, the growing field of ecological digital assessment tools offers valuable insights into monitoring and predicting dementia progression using digital biomarkers in real-world settings. For example, Buegler et al [51] show how these tools provide individualized, context-sensitive data to better understand cognitive performance. Integrating ecological tools with measures like BC-Assess could enhance its utility by capturing real-time data and offering a more comprehensive view of a patient's condition. These tools could also address challenges in tracking cognitive and functional abilities over time, particularly when in-person assessments or informants are unavailable. Further research into combining ecological tools with BC-Assess could refine dementia severity assessments and improve patient outcomes across stages of the disease.

Despite the limitations, this study shows that BC-Assess could be a promising solution for measuring dementia severity. The use of BC-Assess for this purpose will be particularly useful in primary care settings, where DSRS or other comprehensive global assessments may pose implementation challenges. Due to its flexibility, efficiency, and ease of use, BC-Assess can help streamline the assessment process, supporting timely diagnosis and management of dementia. This, in turn, can improve patient outcomes and ease the burden on caregivers.

Acknowledgments

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health (award R44AG078006). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

DH and BH were responsible for the conceptualization and methodology of the study. DH was responsible for the formal analysis and wrote the original draft of the manuscript. All authors reviewed and edited the manuscript.

Conflicts of Interest

All authors report receiving salaries and stock options from BrainCheck.

Multimedia Appendix 1
Raw score (RS) metric and transformed score (TS) calculation for each assessment.
[DOCX File, 7 KB - aging_v8i1e65292_app1.docx]

References

1. Stages of Alzheimer's. Alzheimer's Association. URL: https://alz.org/alzheimers-dementia/stages [accessed 2025-02-18]
2. O'Bryant SE, Waring SC, Cullum CM, et al. Staging dementia using Clinical Dementia Rating Scale Sum of Boxes scores: a Texas Alzheimer's research consortium study. Arch Neurol 2008 Aug;65(8):1091-1095. [doi: 10.1001/archneur.65.8.1091] [Medline: 18695059]
3. Horowitz FD. Developmental theory, prediction, and the developmental equation in follow-up research. In: Developmental Follow-Up: Elsevier; 1994:27-44. [doi: 10.1016/B978-0-12-267855-4.50006-4]
4. Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A. Mapping scores onto stages: mini-mental state examination and clinical dementia rating. Am J Geriatr Psychiatry 2006 Feb;14(2):139-144. [doi: 10.1097/01.JGP.0000192478.82189.a8] [Medline: 16473978]
5. Doody RS, Stevens JC, Beck C. Practice parameter: management of dementia (an evidence-based review). Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001 May 8;56(9):1154-1166. [doi: 10.1212/WNL.56.9.1154] [Medline: 11342679]
6. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
7. Vertesi A, Lever JA, Molloy DW, et al. Standardized Mini-Mental State Examination. Use and interpretation. Can Fam Physician 2001 Oct;47:2018-2023. [Medline: 11723596]
8. Lezak MD, Howieson DB, Bigler ED, Tranel D. Neuropsychological Assessment, 5th edition: Oxford University Press; 2012:1161.
9. Crawford JR. A Handbook of Neuropsychological Assessment: Erlbaum; 1994:454.
10. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437. [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
11. Chan JYC, Yau STY, Kwok TCY, Tsoi KKF. Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: a systematic review. Ageing Res Rev 2021 Dec;72:101506. [doi: 10.1016/j.arr.2021.101506] [Medline: 34744026]
12. Ye S, Sun K, Huynh D, et al. A computerized cognitive test battery for detection of dementia and mild cognitive impairment: instrument validation study. JMIR Aging 2022 Apr 15;5(2):e36825. [doi: 10.2196/36825] [Medline: 35436212]
13. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. J Am Med Assoc 1963 Sep 21;185:914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]
14. Pashmdarfard M, Azad A. Assessment tools to evaluate Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) in older adults: a systematic review. Med J Islam Repub Iran 2020;34:33. [doi: 10.34171/mjiri.34.33] [Medline: 32617272]
15. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. Am J Psychiatry 1982 Sep;139(9):1136-1139. [doi: 10.1176/ajp.139.9.1136] [Medline: 7114305]
16. Sclan SG, Reisberg B. Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4 Suppl 1(3):55-69. [doi: 10.1017/s1041610292001157] [Medline: 1504288]
17. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology (ECronicon) 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
18. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. Int Psychogeriatr 1997;9 Suppl 1(S1):173-176. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
19. Fillenbaum GG, Peterson B, Morris JC. Estimating the validity of the clinical Dementia Rating Scale: the CERAD experience. Consortium to establish a Registry for Alzheimer's Disease. Aging (Milano) 1996 Dec;8(6):379-385. [doi: 10.1007/BF03339599] [Medline: 9061124]
20. Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982 Jun;140(6):566-572. [doi: 10.1192/bjp.140.6.566]
21. Rikkert MGMO, Tona KD, Janssen L, et al. Validity, reliability, and feasibility of clinical staging scales in dementia: a systematic review. Am J Alzheimers Dis Other Demen 2011 Aug;26(5):357-365. [doi: 10.1177/1533317511418954] [Medline: 21914671]
22. Duara R, Loewenstein DA, Greig-Custo MT, et al. Diagnosis and staging of mild cognitive impairment, using a modification of the clinical dementia rating scale: the mCDR. Int J Geriatr Psychiatry 2010 Mar;25(3):282-289. [doi: 10.1002/gps.2334] [Medline: 19565573]
23. Clark CM, Ewbank DC. Performance of the dementia severity rating scale: a caregiver questionnaire for rating severity in Alzheimer disease. Alzheimer Dis Assoc Disord 1996;10(1):31-39. [Medline: 8919494]

24. Cognitive impairment care planning toolkit. Alzheimer's Association. URL: https://www.alz.org/getmedia/d548ca27-309c-43b5-a789-5de837401452/cognitive-impairment-care-planning-toolkit_012623-(1)-(1).pdf [accessed 2025-02-18]
25. Xie SX, Ewbank DC, Chittams J, Karlawish JHT, Arnold SE, Clark CM. Rate of decline in Alzheimer disease measured by a Dementia Severity Rating Scale. Alzheimer Dis Assoc Disord 2009;23(3):268-274. [doi: 10.1097/WAD.0b013e318194a324] [Medline: 19812470]
26. Weinstein JD, Gonzalez ER, Egleton RD, Hunt DA. A paradigm shift for evaluating pharmacotherapy for Alzheimer's disease: the 10-Patient Screening Protocol. Consult Pharm 2013 Jul 1;28(7):443-454. [doi: 10.4140/TCP.n.2013.443]
27. McCulloch K, Collins RL, Maestas KL, LeMaire AW, Pacheco V. Validation and diagnostic utility of the dementia rating scale in a mixed dementia population. Alzheimer Dis Assoc Disord 2014;28(2):168-174. [doi: 10.1097/WAD.0000000000000014] [Medline: 24231456]
28. Reisberg B, Ferris SH, Torossian C, Kluger A, Monteiro I. Pharmacologic treatment of Alzheimer's disease: a methodologic critique based upon current knowledge of symptomatology and relevance for drug trials. Int Psychogeriatr 1992;4 Suppl 1:9-42. [Medline: 1504291]
29. Roalf DR, Moberg PJ, Xie SX, Wolk DA, Moelter ST, Arnold SE. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. Alzheimers Dement 2013 Sep;9(5):529-537. [doi: 10.1016/j.jalz.2012.10.001] [Medline: 23260866]
30. de Craen AJM, Heeren TJ, Gussekloo J. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. Int J Geriatr Psychiatry 2003 Jan;18(1):63-66. [doi: 10.1002/gps.773] [Medline: 12497557]
31. Steffens DC, Otey E, Alexopoulos GS, et al. Perspectives on depression, mild cognitive impairment, and cognitive decline. Arch Gen Psychiatry 2006 Feb;63(2):130-138. [doi: 10.1001/archpsyc.63.2.130] [Medline: 16461855]
32. Sultzer DL, Levin HS, Mahler ME, High WM, Cummings JL. A comparison of psychiatric symptoms in vascular dementia and Alzheimer's disease. Am J Psychiatry 1993 Dec;150(12):1806-1812. [doi: 10.1176/ajp.150.12.1806] [Medline: 8238634]
33. Forsell Y, Winblad B. Major depression in a population of demented and nondemented older people: prevalence and correlates. J Am Geriatr Soc 1998 Jan;46(1):27-30. [doi: 10.1111/j.1532-5415.1998.tb01009.x] [Medline: 9434662]
34. Yang S, Flores B, Magal R, et al. Diagnostic accuracy of tablet-based software for the detection of concussion. Plos One 2017 Jul 7;12(7):e0179352. [doi: 10.1371/journal.pone.0179352]
35. Wallace M, Shelkey M, Hartford Institute for Geriatric Nursing. Katz Index of Independence in Activities of Daily Living (ADL). Urol Nurs 2007 Feb;27(1):93-94. [Medline: 17390935]
36. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990 Dec;46(4):1171-1178. [Medline: 2085632]
37. Zidan M, Arcoverde C, Araújo NBD, et al. Motor and functional changes in different stages of Alzheimer's disease. Arch Clin Psychiatry São Paulo 2012;39(5):161-165. [doi: 10.1590/S0101-60832012000500003]
38. Monaci L, Morris RG. Neuropsychological screening performance and the association with activities of daily living and instrumental activities of daily living in dementia: baseline and 18- to 24-month follow-up. Int J Geriatr Psychiatry 2012 Feb;27(2):197-204. [doi: 10.1002/gps.2709] [Medline: 22223145]
39. Louis ED. Functional correlates of lower cognitive test scores in essential tremor. Mov Disord 2010 Mar 15;25(4):481-485. [doi: 10.1002/mds.22920] [Medline: 20108374]
40. Lee G. Impaired cognitive function is associated with motor function and activities of daily living in mild to moderate Alzheimer's dementia. Curr Alzheimer Res 2020;17(7):680-686. [doi: 10.2174/1567205017666200818193916] [Medline: 32811414]
41. The progression, signs and stages of dementia. Alzheimer's Society. 2021. URL: https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/how-dementia-progresses/progression-stages-dementia [accessed 2025-02-18]
42. Mitchell JC, Dick MB, Wood AE, Tapp AM, Ziegler R. The utility of the Dementia Severity Rating Scale in differentiating mild cognitive impairment and Alzheimer disease from controls. Alzheimer Dis Assoc Disord 2015;29(3):222-228. [doi: 10.1097/WAD.0000000000000057] [Medline: 25187220]
43. Balsis S, Miller TM, Benge JF, Doody RS. Dementia staging across three different methods. Dement Geriatr Cogn Disord 2011;31(5):328-333. [doi: 10.1159/000328116] [Medline: 21555887]
44. Chang YL, Bondi MW, McEvoy LK, et al. Global clinical dementia rating of 0.5 in MCI masks variability related to level of function. Neurology (ECronicon) 2011 Feb 15;76(7):652-659. [doi: 10.1212/WNL.0b013e31820ce6a5] [Medline: 21321338]
45. Bhatt A. Data quality - The foundation of real-world studies. Perspect Clin Res 2023;14(2):92-94. [doi: 10.4103/picr.picr_12_23] [Medline: 37325580]
46. Brorsson B, Asberg KH. Katz index of independence in ADL. Reliability and validity in short-term care. Scand J Rehabil Med 1984;16(3):125-132. [Medline: 6494836]
47. Arik G, Varan HD, Yavuz BB, et al. Validation of Katz index of independence in activities of daily living in Turkish older adults. Arch Gerontol Geriatr 2015;61(3):344-350. [doi: 10.1016/j.archger.2015.08.019] [Medline: 26328478]

48. Kiselica AM, Johnson E, Benge JF. How impaired is too impaired? Exploring futile neuropsychological test patterns as a function of dementia severity and cognitive screening scores. J Neuropsychol 2021 Sep;15(3):410-427. [doi: 10.1111/jnp.12243] [Medline: 33655681]
49. Auer SR, Sclan SG, Yaffee RA, Reisberg B. The neglected half of Alzheimer disease: cognitive and functional concomitants of severe dementia. J Am Geriatr Soc 1994 Dec;42(12):1266-1272. [doi: 10.1111/j.1532-5415.1994.tb06509.x]
50. Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA. Alzheimer disease in the US population: prevalence estimates using the 2000 census. Arch Neurol 2003 Aug;60(8):1119-1122. [doi: 10.1001/archneur.60.8.1119] [Medline: 12925369]
51. Buegler M, Harms R, Balasa M, et al. Digital biomarker-based individualized prognosis for people at risk of dementia. Alzheimers Dement (Amst) 2020;12(1):e12073. [doi: 10.1002/dad2.12073] [Medline: 32832589]

Abbreviations

ADL: activities of daily living
ADL-TS: activities of daily living total score
BC-Assess: BrainCheck standard battery of cognitive assessments
BC-OS: BrainCheck standard battery of cognitive assessments overall score
CDR: Clinical Dementia Rating
DSRS: Dementia Severity Rating Scale
DSRS-TS: Dementia Severity Rating Scale total score
MANOVA: multivariate analysis of variance
ROC: receiver operating characteristic
ROC-AUC: area under the receiver operating characteristic curve

Edited by PH Chau; submitted 13.08.24; peer-reviewed by A Vella, M Colombo; revised version received 19.12.24; accepted 25.12.24; published 26.02.25.

Please cite as:
Huynh D, Sun K, Patterson M, Hosseini Ghomi R, Huang B
Performance of a Digital Cognitive Assessment in Predicting Dementia Stages Delineated by the Dementia Severity Rating Scale: Retrospective Study
JMIR Aging 2025;8:e65292
URL: https://aging.jmir.org/2025/1/e65292
doi:10.2196/65292

Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study

Abeer Badawi[1*], PhD; Somayya Elmoghazy[1*], MSc; Samira Choudhury[2,3], MSc, MD; Sara Elgazzar[4], BSc; Khalid Elgazzar[1,5], PhD; Amer M Burhan[2,6], MBChB, MSc

[1]Internet of Things (IoT) Research Lab, Department of Electrical and Computer Engineering, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

[2]Ontario Shores Centre for Mental Health Sciences, Whitby, ON, Canada

[3]Department of Psychiatry, Western University, London, ON, Canada

[4]Faculty of Science, Ontario Tech University, Oshawa, ON, Canada

[5]Canadian University of Dubai, Dubai, United Arab Emirates

[6]Temerty Faculty of Medicine, University of Toronto, Toronto, Canada

[*]these authors contributed equally

Corresponding Author:
Somayya Elmoghazy, MSc
Internet of Things (IoT) Research Lab, Department of Electrical and Computer Engineering, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

Abstract

Background: Dementia is a progressive neurodegenerative condition that affects millions worldwide, often accompanied by agitation and aggression (AA), which contribute to patient distress and increased health care burden. Existing assessment methods for AA rely heavily on caregiver reporting, introducing subjectivity and inconsistency.

Objective: This study proposes a novel, multimodal system for predicting AA episodes in individuals with severe dementia, integrating wearable sensor data and privacy-preserving video analytics.

Methods: A pilot study involving 10 participants was conducted at Ontario Shores Mental Health Institute. The system combines digital biomarkers collected from the EmbracePlus (Empatica Inc) wristband with video-based behavioral monitoring. Facial features in video frames were anonymized using a masking tool, and a deep learning model was used for AA detection. To determine optimal performance, various machine learning and deep learning models were evaluated for both wearable and video data streams.

Results: The Extra Trees model achieved up to 99% accuracy for personalized wristband data, while the multilayer perceptron model performed best in general models with 98% accuracy. For video analysis, the gated recurrent unit model achieved 95% accuracy and 99% area under the curve, and the long short-term memory model demonstrated superior response time for real-time use. Importantly, the system predicted AA episodes at least 6 minutes in advance in all participants based on wearable data.

Conclusions: The findings demonstrate the system's potential to autonomously and accurately detect and predict AA events in real-time. This approach represents a significant advancement in the proactive management of behavioral symptoms in dementia care.

(JMIR Aging 2025;8:e68156) doi:10.2196/68156

KEYWORDS

dementia; agitation; artificial intelligence; wearable sensors; video detection; multimodal sensing

Introduction

Background

Dementia is a neurodegenerative condition that leads to a progressive decline in cognition and is one of the leading causes of death, disability, and hospitalization in Canada and worldwide. Currently, dementia is the seventh cause of death worldwide [1]. Worldwide, over 55 million individuals are living with dementia; as the ratio of older people increases, this number will grow to 78 million by 2030 and 139 million by 2050, making dementia a major global health crisis [1]. In addition to cognitive and functional decline, people living with dementia also experience noncognitive neuropsychiatric symptoms (NPS) during their illness [2]. NPS commonly includes agitation, aggression, apathy, symptoms of psychosis, delusions, hallucinations, and disturbances of sleep and appetite. Among NPS, agitation and aggression (AA) occur frequently

in severe cases and are a common source of distress for patients and caregivers [3]. They commonly occur during care and are believed to be manifestations of perceived or real unmet needs [3]. Behaviors of AA include pacing, rocking, gesturing, restlessness, shouting, throwing objects, and destroying property [4]. These symptoms are the leading cause of hospitalizations, extended length of stay of inpatients, and increased demand for placement in long-term care facilities [5]. AA enormously burdens people living with dementia, their families, caregivers, and health care systems.

In current practices, AA is commonly assessed through caregiver reports. Many observational methods have been developed, including the Neuropsychiatric Inventory [6] and the Cohen-Mansfield Agitation Inventory [7]. These assessments are based on manual observations, which are subject to potential bias depending on the caregiver's memory or emotional state. It is possible to address these limitations by using artificial intelligence (AI) and predictive algorithms to predict episodes of AA in people living with dementia. In 2025, AI technologies are expected to be worth an estimated US $36 billion [8]. There is growing evidence that combining AI and sensory technologies to develop a solution for NPS detection will guide the provision of personalized interventions for people living with dementia [9-11]. The timely detection of critical events in people living with dementia using digital technologies is gaining wide acceptance [12,13]. Predicting and managing AA in people living with dementia requires innovative approaches that integrate multiple data sources. Wearable sensors and video-based monitoring systems provide unique opportunities for the real-time detection of AA and preagitation behaviors, but challenges such as privacy concerns and scalability have limited their adoption in clinical settings. This study addresses these challenges by combining biometric data from wearable sensors with AI-driven video analysis, enabling real-time detection and prediction of AA. This integrated approach aims to facilitate timely interventions, reduce care costs, and improve outcomes for both patients and caregivers. We conducted a pilot study in the Geriatric Dementia Unit (GDU) and the Geriatric Transitional Unit at the Ontario Shores Center for Mental Health Sciences [14]. We combine biometric data from the EmbracePlus wristband [15] and video data from CCTV cameras installed in common areas in both units. Machine learning and deep learning techniques, including Extra Trees, Gradient Boosting, Random Forest, multilayer perceptron (MLP), and recurrent neural networks (RNNs), are used to analyze biometric and skeletal data extracted from both the wristband and the video cameras. We achieved high accuracy in detecting AA from the EmbracePlus wristband through comprehensive data preprocessing, feature extraction, the Extra Trees classification algorithm for personalized models, and the MLP algorithm for the general model. Additionally, AA detection was enhanced by analyzing real-time video feeds with skeletal key points and using RNN-based neural networks, particularly long short-term memory (LSTM) and gated recurrent unit (GRU) [16]. These networks, optimized for real-time processing, facilitate timely interventions. The pilot study demonstrated the system's effectiveness through both the wristband and video detection.

Related Work

The growing number of people living with dementia causes significant challenges for health care systems and caregivers. One of these challenges is to deal with symptoms of AA that increase with the severity of dementia. Recent advancements in multimodal sensing technologies, including cameras and wearable wristbands, have shown promise in monitoring and managing AA in people living with dementia [10,17,18]. Wearable devices, capable of capturing physiological signals such as acceleration, heart rate, and skin conductance, have demonstrated potential for real-time AA detection [19-21]. Recent advancements in wearable sensor technology and AI have shown promise in addressing the early detection of AA behavior when focusing on signal processing and machine learning to extract features and classify AA events [22].

However, real-time video-based monitoring systems to monitor AA behavior in dementia patients are an area of interest for researchers today [10,17,23]. These systems use cameras positioned in patients' rooms or common areas to consistently record and monitor patients' behavioral data. Some research [24,25] used video cameras to detect AA from previously recorded videos at the Specialized Dementia Unit, Toronto Rehabilitation Institute, Canada. Their system focused on offline AA detection. Another work collected the training dataset from healthy people who imitated agitated hand movements [26]. Researchers have highlighted the importance of real-time feedback, which allows for timely interventions and reduces the severity and duration of AA events. This can help improve the quality of life and reduce the stress caused by agitated behaviors in both patients and caregivers. Researchers have also considered privacy factors and concerns while ensuring the accuracy of these systems. The work done in studies by Mishra et al [24] and Marshall [26] has proposed different masking methods for the people in the video frames that allow for AA detection while keeping the patients' identities and features hidden.

Multimodal sensing, which combines data from multiple sensors, types, and sources, has emerged as a promising approach to understanding AA in people living with dementia [27,28]. It enhances the detection of early signs of AA and the identification of relevant triggers [23,29]. Clinical trials evaluating the use of multimodal sensing in the context of dementia and AA behavior are crucial in this type of research. These trials assess the feasibility and efficacy of monitoring systems that combine various data sources to inform clinical decision-making. The results of these trials will provide valuable insights into the practical implications of multimodal sensing in real-world health care settings. Most existing studies on video systems and wearable sensors for AA detection in real-time have been conducted in controlled laboratory settings or residential care facilities with limited datasets [18,30]. Research is needed to use more diverse and general datasets gathered and applied in real hospital settings. This is essential since the data from hospitals would be more representative of people with severe dementia who are more prone to AA. Research in such a real-world setting with data gathered from real patients is vital for developing realistic and applicable solutions and ensuring

effective treatment and care for people living with dementia [18,30].

Moreover, it is challenging to identify the actual start and end times of AA episodes. This is because the AA events are extracted from nurse notes, which are not accurate and prone to human or individual error. Many AA episodes may also be overlooked and mislabeled as nonagitation. In addition to the datasets, there is a high demand for accurate and reliable end-to-end real-time monitoring solutions to actively predict and respond to AA events. There is also still a need for an in-depth investigation of AI and its features in addition to an in-depth investigation of preagitation patterns and signs that could trigger AA in people living with dementia. These investigations are necessary to indicate the usefulness of preagitation signs and patterns in early prediction. Digital biomarkers can help detect and predict AA early in real time [31]. The investigation by Alam et al [32] shows the correlation between motion biomarkers collected from the accelerometer and the early AA signs, which is particularly useful for personalized AA detection models.

This paper presents a unique system that combines video analysis and wearable sensor data to predict AA in people living with dementia. The video feeds are crucial for identifying the precise start and end of each AA episode. This precise timing enables us to incorporate data analysis from both the wristband and video footage into our research. We focus on AI and advanced feature engineering to improve the detection accuracy of AA in people living with dementia. We significantly enhance our chances of detecting AA by using 2 distinct yet cooperative models (one based on physiological data and the other on video analysis). This integrated approach also opens the doors to incorporating additional predictive methods, such as audio-based AA detection. The combined use of video-based systems and wearable wristbands offers deeper insights into AA management. While notable advancements have been achieved, further enhancements are needed, especially in real-time accuracy and system applicability. We are confident that our current research explores a vital area and will contribute new knowledge to the field and lay a solid foundation for future advancements.

Methods

Ethical Considerations

The research commenced in 2019. This study was approved by the Joint Research Ethics Board (JREB) at Ontario Shores Centre for Mental Health Sciences and Ontario Tech University (JREB Number: 21-011-B). Informed consent was obtained from all participants' substitute decision makers, as the participants had advanced dementia and were unable to provide consent themselves. Data collected by the cameras were masked, with any identifiable objects in the video frames blurred (a video demonstration was provided to the research ethics board [REB] chair and privacy officer). All information obtained from participants was kept confidential. Computer-based data were stored in password-protected databases, and paper-based files were kept in locked cabinets. Access to all data were restricted to authorized study personnel, who followed the confidentiality regulations of the JREB. Participants were not compensated for their participation.

This study was approved by the JREB at Ontario Shores Centre for Mental Health Sciences and Ontario Tech University (JREB Number: 21-011-B). Informed consent was obtained from all participants' substitute decision-makers, as the participants had advanced dementia and were unable to provide consent themselves. Data collected by the cameras were masked, with any identifiable objects in the video frames blurred (a video demonstration was provided to the REB chair and privacy officer). All information obtained from participants was kept confidential. Computer-based data were stored in password-protected databases, and paper-based files were kept in locked cabinets. Access to all data was restricted to authorized study personnel, who followed the confidentiality regulations of the JREB. Participants were not compensated for their participation.

Study Design

A significant challenge was presented using video cameras to document the behaviors and activities of patients, staff, and visitors in the public spaces of hospital inpatient units. Consent was secured for these patients through substitute decision makers considering their advanced dementia condition. At admission, capacity assessments were performed, and substitute decision makers were approached for permission to involve patients in potential research. The recording of video cameras was restricted to the common areas where patients usually gather during the day, and audio capture was turned off throughout the data collection period.

This study collected participant data using an EmbracePlus wristband [15] and video cameras installed in the GDU and the Geriatric Transitional Unit at the Ontario Shores Center for Mental Health Sciences. A total of 10 participants were recruited based on inclusion criteria, including being aged 60 years or older, a diagnosis of moderate to severe major neurocognitive disorder as determined by the Mini-Mental State Examination [33], and the ability to ambulate independently with or without a walking aid. Additionally, participants had to meet the agitation criteria defined by the Agitation Definition Working Group of the International Psychogeriatric Association [34], with a Functional Assessment Staging Tool scale score between 6a and 6e [35]. Moreover, each unit was equipped with a single AXIS M3077-PLVE Network Camera [36], and one AXIS P3225-VE Mk II camera [37] was installed in the hallway of the GDU to capture relevant footage. Participants wore the EmbracePlus wristband for 24 to 72 hours on 3 separate occasions within a 6-week study period. The wristband collected physiological parameters such as heart rate, electrodermal activity, and skin temperature.

During data collection (Multimedia Appendix 1), clinical staff, who are part of the research team of this study, monitored participants for episodes of AA, noting the start and end times of each event. CCTV cameras recorded footage to provide precise timestamps and additional context, as there can be a delay between observed behavior and recorded notes. To maintain privacy, faces and identifiable features in the video footage were blurred. Once an AA event is identified by the

staff in the video, the skeletal points of the recruited participants are extracted and added to the dataset, and all other skeletal points are discarded. The skeletal key points were analyzed using deep learning models to classify episodes of AA. Data from wearable sensors and clinical notes were integrated and analyzed to identify physiological and behavioral patterns preceding episodes of AA. The collected data formed the basis for developing and evaluating our multimodal system for real-time AA detection. The results demonstrate the system's feasibility and effectiveness in detecting and analyzing AA episodes.

Event Classification

The proposed system collects the biomarkers using the EmbracePlus wristband, which is considered the state-of-the-art wearable device for continuous health monitoring in the market today [15]. The device combines digital biomarkers, robust sensors, and a user-friendly design to continuously monitor participants with various health conditions. It collects electrodermal data of detected slight changes in skin conductance from the skin surface, and the data of a photoplethysmogram that calculates the pulse rate and pulse rate variability measurements, skin temperature, and raw accelerometry data for motion detection. The collected signals are sent to the cloud-based EmbracePlus Care platform [15] through a Bluetooth-connected gateway (eg, a smartphone).

The first type of data we deal with is the raw data from the accelerometer, heart rate, temperature, and electrodermal signals. We follow several preprocessing steps to clean, filter, and apply 1-minute window segmentation for the raw signals [21,38]. We then extract features from the signals, as shown in our previous work, from the statistical, time domain, frequency domain, and time-frequency domain with around 150 features [21,38]. Lastly, we evaluate multiple classification techniques, namely Random Forest, Extra Trees, and Gradient Boosting, to classify AA events. The performance of each model is evaluated using standard classification metrics such as accuracy, precision, recall, area under the curve (AUC), and F_1-score. Furthermore, we collect the digital biomarkers that are preprocessed data derived from Empatica's algorithms and calculate them minute-by-minute. The second type is the digital biomarkers, which include pulse rate variability, respiratory rate, movement intensity, accelerometer magnitude SD, steps, skin conductance level, wearing detection, temperature, and sleep detection as shown in Table 1. Digital biomarkers can effectively and accurately oversee human health from a distance, consistently, and without causing disruption. This applies across a spectrum of health conditions [15]. Figure 1 shows the classification workflow from the EmbracePlus wristband using raw data and digital biomarkers.

Table . The digital biomarker data description from the EmbracePlus wristband.

Digital biomarkers	Definition
PR[a]	The algorithm uses the photoplethysmogram and accelerometer data for PR monitoring with estimates on 10-second windows.
PRV[b]	The algorithm analyzes the photoplethysmogram for intermittent PRV, using accelerometer signals with nonoverlapping windows.
RR[c]	The algorithm processes the photoplethysmogram and accelerometer data to calculate RR values expressed in breaths per minute.
Movement intensity	The algorithm calculates activity count, steps, and accelerometer SD from the accelerometer sensor.
SCL[d]	SCL estimation from EmbracePlus electrodermal activity signal, output every 1 minute with nonoverlapping windows.
Wearing detection	The algorithm correlates device status with the photoplethysmogram patterns, indicating wearing time.
Temperature	The algorithm analyzes EmbracePlus data for continuous peripheral temperature estimation.

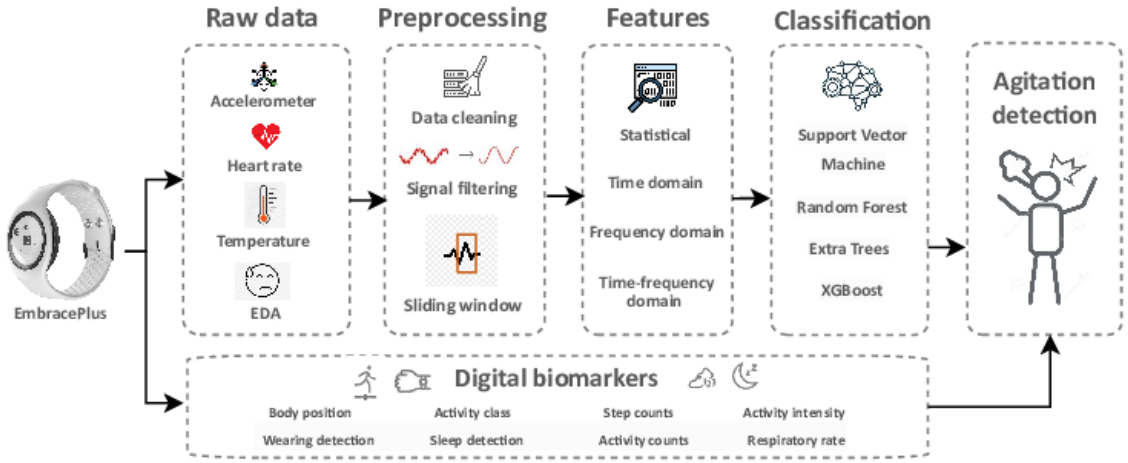
[a]PR: pulse rate.

[b]PRV: pulse rate variability.

[c]RR: respiratory rate.

[d]SCL: skin conductance level.

Figure 1. The workflow of the proposed classification system using the EmbracePlus wristband. EDA: electrodermal activity.



The proposed work focuses on investigating the data from people living with dementia using machine learning, 2 of which were thoroughly investigated in our previous work [21,38]. The results concluded the most important features for this problem after performing feature engineering and proved that personalized models on individual patients outperform generic models. In this work, we report the results of the personalized model on 10 different participants from the Ontario Shores Mental Health Institute. We test our system in real-time once we determine the optimal classification system to predict AA events. In the real-time detection phase, real-time raw data is transmitted from the wristband. Following this, features are extracted from each 1-minute window, and these specific features are fed into the customized model to classify whether the data is considered normal or indicative of AA. The outcomes are subsequently transmitted to the backend system, and the health care provider is notified if the patient is agitated.

Video-Based Analysis

In addition to collecting data on physiological biomarkers, this study incorporates video analysis data for AA prediction. This approach uses an extra cooperative model, which improves our overall AA detection system and allows us to get the precise duration, including start and end times, of the collected AA episodes. Moreover, once an AA episode is detected, the cameras record a previously set preagitation, making it easier to observe any visual preagitation signs. We aim to provide real-time alerts to health care providers for timely intervention. The setup includes 3 CCTV cameras installed and a PC in the attending psychiatric office with access to this footage. Our system operates in 2 phases: the offline phase for manual labeling and model training, and the real-time stage for running the model. To protect the privacy of the participants and the staff present, we blur all faces and run OpenPose (Carnegie Mellon University) to capture the movement data of the participants [39-42]. OpenPose is an advanced real-time system for multiperson 2D pose estimation. It also helps to anonymize individuals in video frames. OpenPose uses convolutional neural networks to detect human body parts and map their skeletal structure onto the image or video frame. This allows for a detailed representation of movement data present in the collected frames. The model is trained on features extracted from skeletal point coordinates instead of the raw video frames. This approach has been recently used by researchers for AA detection in people living with dementia and has proven to be as successful in detecting AA while preserving the privacy of the people present [24,26].

After this, we use a preprocessing phase to enhance the generalizability of the model across various environments and datasets. This phase involves the elimination of extraneous noise that could otherwise impede model performance. The model considers the variations in camera angles and participant positioning within the frame, which can significantly influence the coordinate data. We calculate Euclidean distances and angle measurements between specific skeletal coordinates to determine movements. For example, the measured distance between the torso and feet is useful to identify potential kicking actions, which may indicate AA in certain contexts. Table 2 summarizes all 47 features extracted from these distance and angle measurements. Before training, a feature analysis step is introduced to remove highly correlated features and reduce the dimensionality of the model. This process reduced the features to 39 features. We tested the system on the same 10 participants whose wearable sensor data was used earlier.

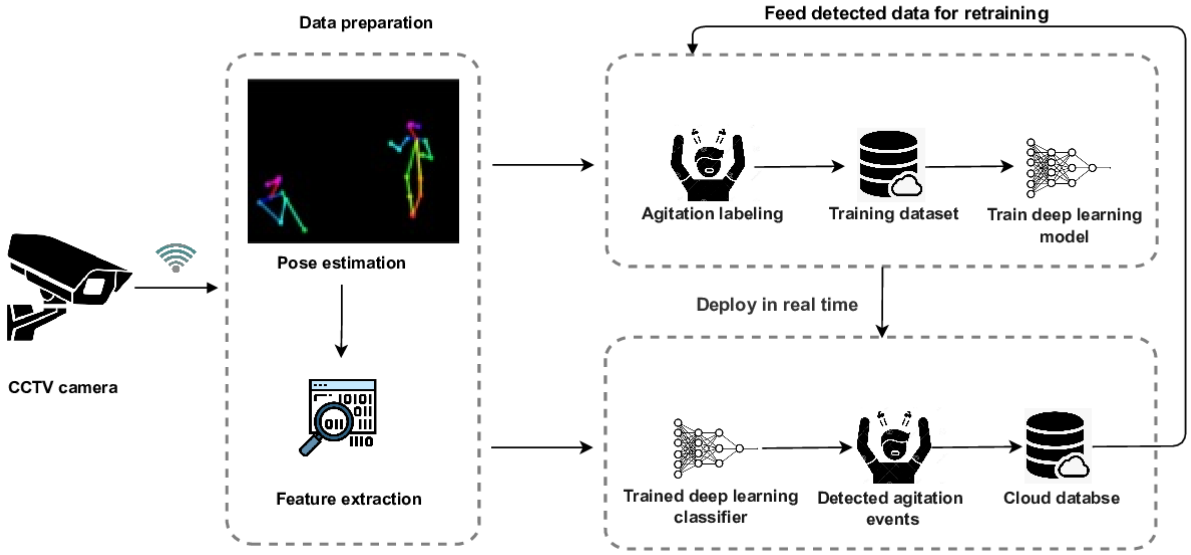
Table . Description of the extracted features from the skeletal data.

Feature	Description
eu_1-eu_14	Euclidean distances between different key point pairs; eu 1 represents the Euclidean distance between key point 1 and the previous position of key point 1
eu_1_3-eu_1_14	Euclidean distances between key point 1 and various other key points
por_2_1-por_14_1	POR[a] values between key points 2 - 14 and key point 1
ang_1_2-ang_1_14	Angles between key point 1 and key points 2 - 14

[a]POR: point of reference.

The offline and real-time (online) stages of the system are detailed in Figure 2. In the offline stage, we preprocess and extract features from skeletal data, which are then labeled using nurse notes from patient medical records. Access to the collected videos is restricted to the computer in the psychiatric office, and they are retained only until the AA episodes are accurately labeled with their start and end times. Once the dataset is finalized, all videos are securely discarded. Using this dataset, we train a deep learning model to differentiate between AA and nonagitation events. Our focus is on capturing a range of behaviors, from violent or aggressive actions to repetitive motions such as pacing or chair rocking. Hence, we use models that consider sequences of actions, such as RNN-based neural networks, to effectively recognize these sequences of actions. We specifically use and compare the results of the GRU and LSTM models. Both are adept at analyzing sequences of actions. The LSTM model is designed to capture long-term dependencies within sequences. The sophisticated cell structure of LSTM cells makes it highly effective in maintaining context over long intervals. However, the GRU model uses a simpler architecture that aims to achieve results comparable to LSTM models but with lower computational costs. Both models used in our research are composed of a single LSTM or GRU cell, followed by a fully connected sigmoid layer for the binary classification of AA episodes.

Figure 2. The system architecture of the video-based detection system.



In the real-time stage, we deploy our offline-trained model to classify AA in real-time. The model first processes real-time video data from hospital cameras. Video frames are analyzed using OpenPose, extracting features similar to those in the offline stage. We use a fixed-size window to input frame sequences into the classifier. Upon detecting AA, the system records the event, including a 5-minute buffer before and after the incident to capture the entire context. This approach helps identify potential triggers and patterns that lead to AA. The psychiatrist reviews these detected events for accuracy and confirms AA events, which are then added to our training dataset with appropriate labels. The model is then retrained in the offline stage with the added data. The primary goal of retraining the model is to continuously adapt and improve the model with newly detected data.

Results

Overview

A pilot study was conducted to validate the effectiveness and feasibility of the proposed system at Ontario Shores Mental Health Hospital. This initial investigation aimed to provide valuable insights into the system's functionality, usability, and overall potential before the implementation on a larger population. Details of the enrolled participants and the data

collected can be found in Table 3. Upon enrollment, the participants wore the EmbracePlus wristband for 24 to 72 hours. We turned on the cameras installed in the unit during the data collection days to record the participants' activities. Lastly, we assigned a nurse to observe the participant and provide a detailed report of behavior, AA events, and any abnormal behavior. During the 3 days, we collected 6 AA events, ranging from 2 to 23 minutes per AA event, with a total of 20 - 32 minutes of AA labels and 560 - 581 minutes of normal labels for each participant. The following sections will present the results from the EmbracePlus wristband and video cameras in detail.

Table . Overview of the demographic of the participants and the total collected data.

Participant	Gender	Age	Recruitment date	Total collected data (hours)
1	Female	83	August 2023	54.21
2	Female	63	January 2024	48.13
3	Female	77	March 2024	45.3
4	Female	78	March 2024	96.4
5	Female	79	March 2024	99.4
6	Male	79	May 2024	46.7
7	Male	67	August 2024	36
8	Male	68	August 2024	35.5
9	Male	85	September 2024	95.5
10	Male	86	September 2024	54.6

Performance Evaluation

The EmbracePlus Wristband Raw Data

This study uses raw data obtained from 4 wristband signals and personalized models, which achieved superior accuracy in AA detection from people living with dementia in previous research [21,38]. Subsequently, we conducted a comparative analysis of multiple machine learning and deep learning algorithms for AA detection, including Random Forest, Extra Trees, Gradient Boosting, and MLP. We trained and tested a personalized model for every participant and reported the evaluation results in Table 4. The dataset for each participant was randomly split into 70% for training and 30% for testing.

Table . Comparative performance metrics using raw data.

Participant	Extra Trees			Gradient Boosting			Random Forest			MLP[a]		
	Acc[b]	AUC[c]	Recall	Acc	AUC	Recall	Acc	AUC	Recall	Acc	AUC	Recall
1	0.98[d]	0.99[d]	0.99[d]	0.97	0.98	0.98	0.98[d]	0.99[d]	0.99[d]	0.97	0.98	0.98
2	0.90[d]	0.98[d]	0.96	0.88	0.96	0.93	0.88	0.96	0.94	0.88	0.96	0.97[d]
3	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.99	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
4	0.91[d]	0.99[d]	0.97[d]	0.89	0.97	0.96	0.88	0.97	0.99	0.89	0.96	0.98
5	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.95	0.90	0.96
6	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
7	0.92[d]	0.98[d]	0.97	0.89	0.96	0.96	0.89	0.97	0.96	0.90	0.96	0.99[d]
8	0.91[d]	0.98[d]	0.97	0.91	0.96	0.96	0.88	0.97	0.95	0.89	0.96	0.98
9	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.98	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
10	0.99[d]	0.99[d]	0.99[d]	0.98	0.99[d]	0.99	0.98	0.99[d]	0.98	0.99[d]	0.99[d]	0.99[d]
All	0.85	0.94	0.93	0.98	0.96	0.90	0.93	0.95	0.65	0.99[d]	0.98[d]	0.98[d]

[a]MLP: multilayer perceptron.

[b]Acc: accuracy.

[c]AUC: area under the curve.

[d]Bold text indicates the best results for each participant in each metric.

The Extra Trees model outperformed the rest of the models for most of the participants, followed by the MLP model, which achieved similar results for 4 of the 10 participants. For example, the Extra Trees model achieved an accuracy of 98.67%, an AUC of 99.1%, and a recall of 99.76% for participant #1. It achieved the highest accuracy and AUC for participant #2 of 90% and 98%, respectively. For participants, #3, #6, #9, and #10, both the Extra Trees and MLP models achieved 99% across all evaluation metrics. These results underscore the efficacy of the chosen features, preprocessing methodologies, and up-sampling techniques. When tested on all 10 participants together, the Extra Trees model achieved a lower accuracy of 85%. The Gradient Boosting and Random Forest achieved higher accuracies in comparison with 98% and 93%, respectively. The Random Forest model, however, performed very poorly in other evaluation matrices, with a recall of 67% for the general model. The MLP model achieved a slightly higher accuracy than Gradient Boosting of 99% and even a higher AUC and recall of 98%, performing very similarly to the personalized models. This highlights the potential for a general model when enough data are collected.

Table 5 presents a summary of the statistical analysis of model performance metrics across all models. As expected, Extra Trees is the top-performing model for participant-level predictions, with the highest mean accuracy of 0.95 and a 95% CI of 0.928-0.986. This indicates consistent performance across the evaluation set. Pairwise statistical tests further confirmed the superior accuracy of Extra Trees compared to Gradient Boosting (P=.001), Random Forest (P=.006), and MLP (P=.002). Although MLP demonstrated a comparable mean accuracy of 0.94 and even outperformed Extra Trees in the general model setting, it exhibited a broader CI (0.909-0.979) and a similar SD (0.04). In terms of other metrics, Extra Trees also achieved the highest mean AUC of 0.98 and a narrow CI of 0.970-0.990. MLP and Gradient Boosting delivered competitive results in AUC 0.97 and recall 0.96. These findings suggest that while MLP performed well in the general model, Extra Trees consistently outperformed other models in the personalized model evaluations. This consistency identifies Extra Trees as the most reliable model for this multimodal system. Additionally, the observed trends highlight the importance of focusing on personalized models to consider individual variations.

Table . Statistical analysis of model performance metrics across all models.

Metric	Extra Trees	Gradient Boosting	Random Forest	MLP[a]
Mean accuracy	**0.95**[b]	0.94	0.94	0.94
SD (Acc[c])	0.04	0.04	0.05	0.04
95% CI (Acc)	0.928-0.986	0.912-0.976	0.905-0.977	0.909-0.979
Paired P value (vs Extra Trees)	—[d]	0.001	0.006	0.002
Mean AUC[e]	**0.98**[b]	0.97	0.96	0.97
SD (AUC)	0.01	0.01	0.01	0.01
95% CI (AUC)	0.970-0.990	0.960-0.980	0.950-0.970	0.960-0.980

[a]MLP: multilayer perceptron.

[b]Bold text indicates the best result for each metric.

[c]Acc: accuracy.

[d]Not applicable.

[e]AUC: area under the curve.

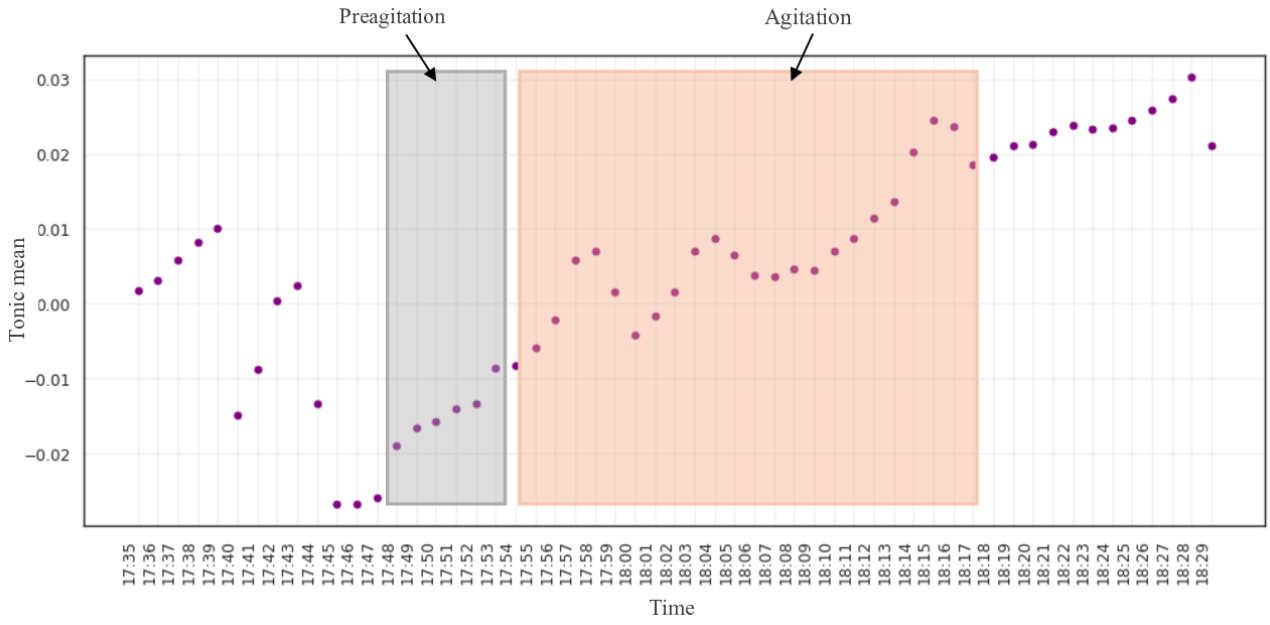
Furthermore, we studied the importance of different features for every participant (Multimedia Appendix 2). For participant #1, the top 10 features contributing to accurate AA classification using the Extra Trees model revealed that 5 were from electrodermal activity, 3 from the accelerometer, 1 from heart rate, and 1 from temperature. Notably, the electrodermal tonic mean was the most critical feature, suggesting a strong link between AA episodes and fluctuations in electrodermal activity, which is commonly associated with emotional arousal. For participant #2, the most important features were primarily the accelerometer and temperature signals. Features related to energy, root mean square, and variability in acceleration played a key role in AA classification. Additionally, temperature fluctuations were also significant, suggesting that both movement patterns and body temperature changes could indicate agitation onset in this participant.

For participant #3, the temperature-related features were the most dominant in identifying AA episodes. The maximum temperature value, temperature root mean square, and energy were among the top contributors, highlighting a strong correlation between changes in body temperature and agitation. These findings emphasize the individual variability in AA predictors, reinforcing the necessity of personalized models for accurate classification. While some participants exhibit agitation-related physiological changes in electrodermal activity, others may show significant patterns in movement or temperature variations, underscoring the importance of a multimodal feature selection approach.

As the electrodermal tonic mean was the top feature to classify AA for participant #1, we investigated the AA labels. Figure 3 shows the tonic mean values of participant #1 from the electrodermal signal during labeled AA events from the camera

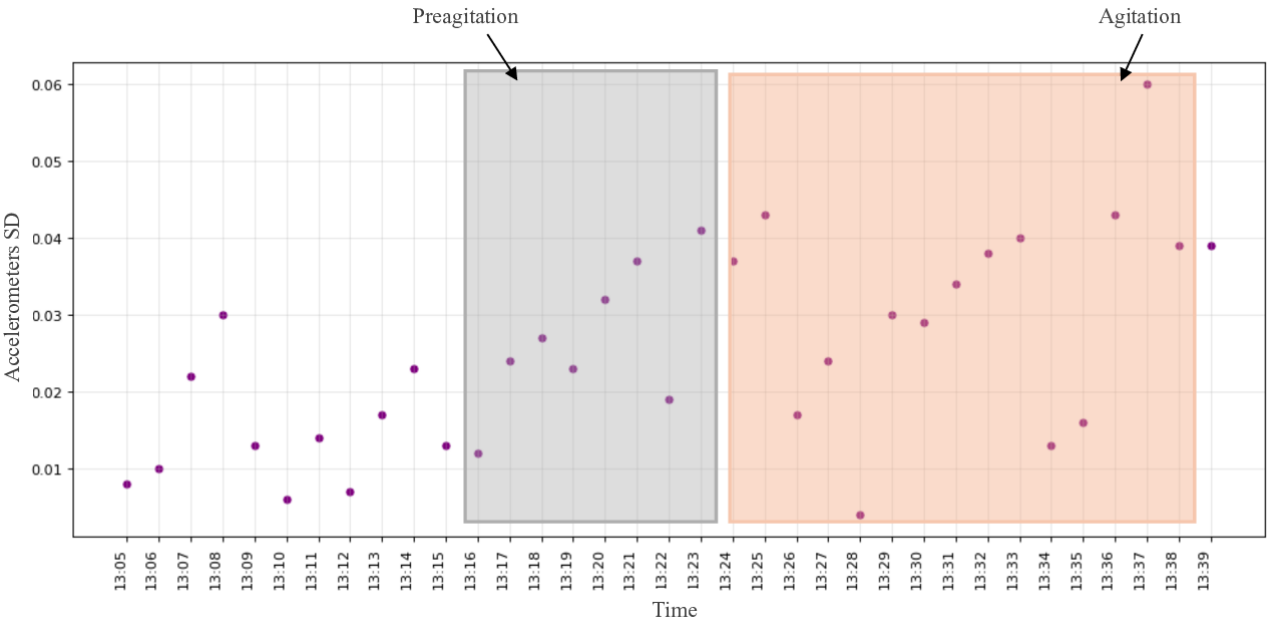
and nurse notes (highlighted in red). This event occurred during the second day and lasted for 23 minutes from 5:55 PM to 6:17 PM. This observation suggests that the patient's AA was related to the electrodermal signal connected to the emotions. We also observed an apparent change to the data before the actual AA occurred, which we manually marked as preagitation labels (highlighted in blue).

Figure 3. The EmbracePlus wristband raw data with agitation and preagitation annotation: the tonic mean plot for participant #1.



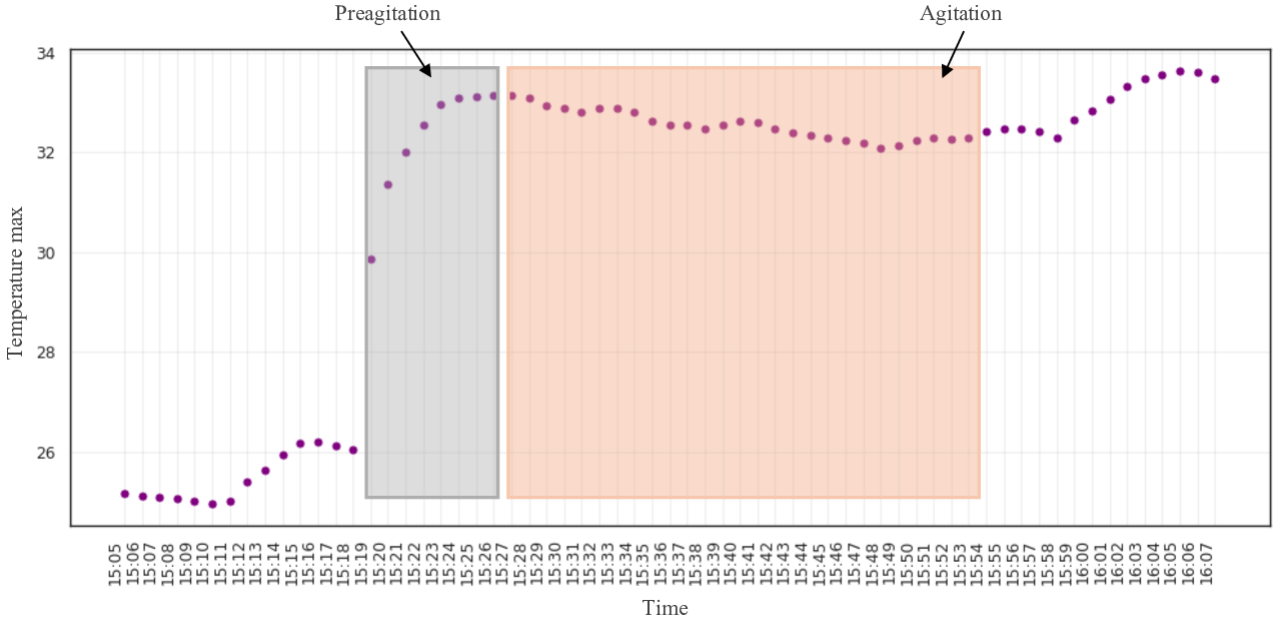
For participant #2, the acceleration features were the top features in identifying the AA event. Figure 4 shows one of the AA events for this participant occurring between 1:24 PM and 1:38 PM using the accelerometer data. Moreover, a change in the pattern of the signal 8 minutes before the event is manually labeled as preagitation.

Figure 4. The EmbracePlus wristband raw data with agitation and preagitation annotation: the accelerometer plot for participant #2.



For participant #3, the most dominant features were the temperature features, so an example of the temperature readings for an AA event is shown in Figure 5. Just as before, a change in the pattern before the observed AA event is manually labeled as preagitation. These observations suggest the potential for detecting preagitation patterns from raw data, enabling the prediction of AA before it occurs.

Figure 5. The EmbracePlus wristband raw data with agitation and preagitation annotation: the temperature plot for participant #3. max: maximum.



The Wristband EmbracePlus Digital Biomarkers

We explored all the digital biomarkers offered by EmbracePlus. Taking the first 3 participants as an example, we observed that pulse rate, activity counts, and activity class were the leading indicators for AA detection for the first 3 participants. In Figures 6-8 , the same AA events discussed in the previous subsection for the 3 participants are illustrated, and the values during labeled AA events from the camera and nurse notes are highlighted in red. Additionally, we observed a noticeable change in the data before the onset of AA, manually designated as preagitation labels and highlighted in gray. The manual labeling of the preagitation was done after reviewing all the signals for the participants and noting the same change across multiple patterns.

Figure 6. The activity counts from the digital biomarkers data for participant #2.

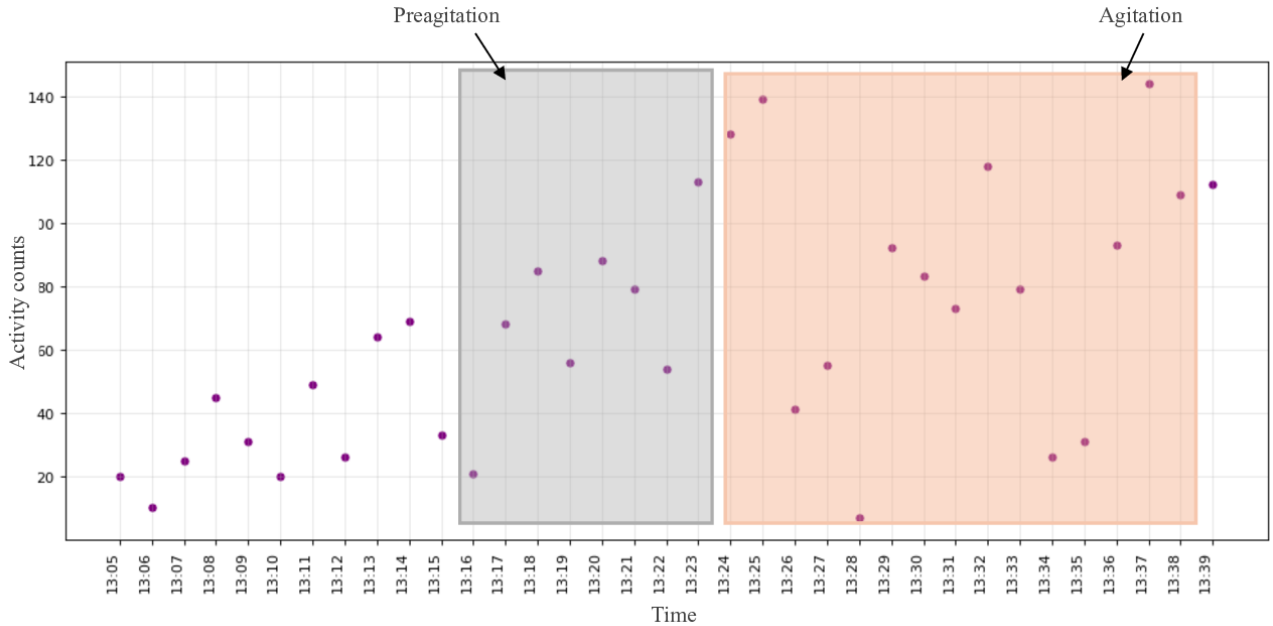


Figure 7. The pulse rate from the digital biomarkers data for participant #3.

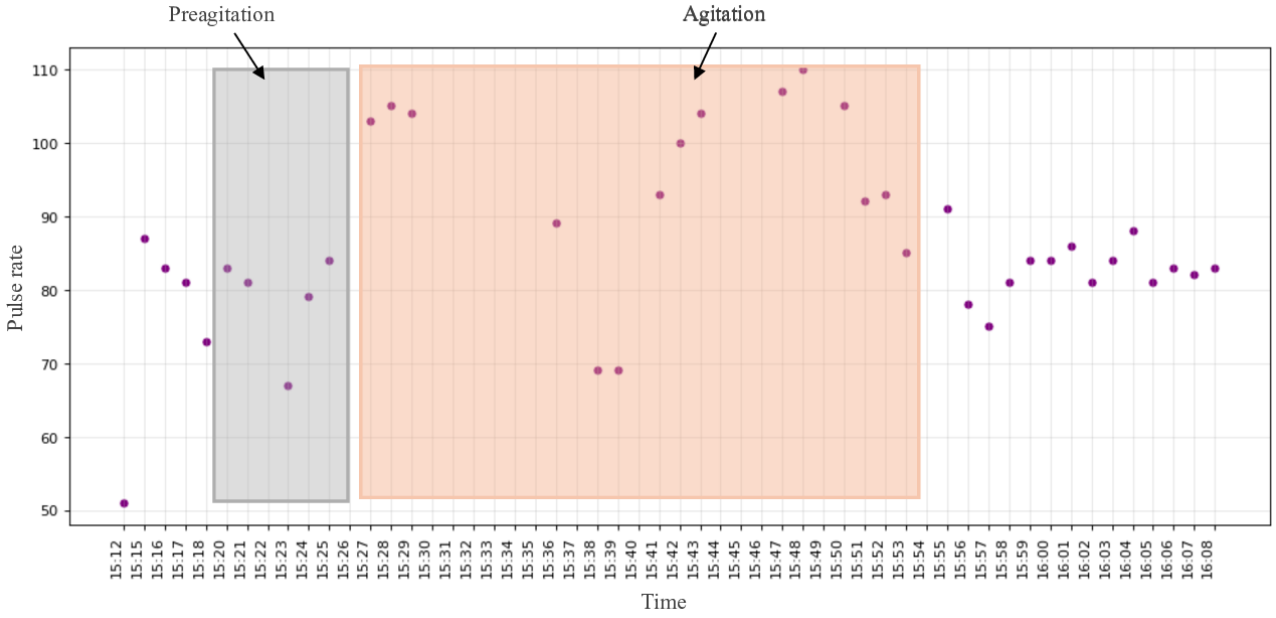


Figure 8. The activity class from the digital biomarkers data for participant #1.

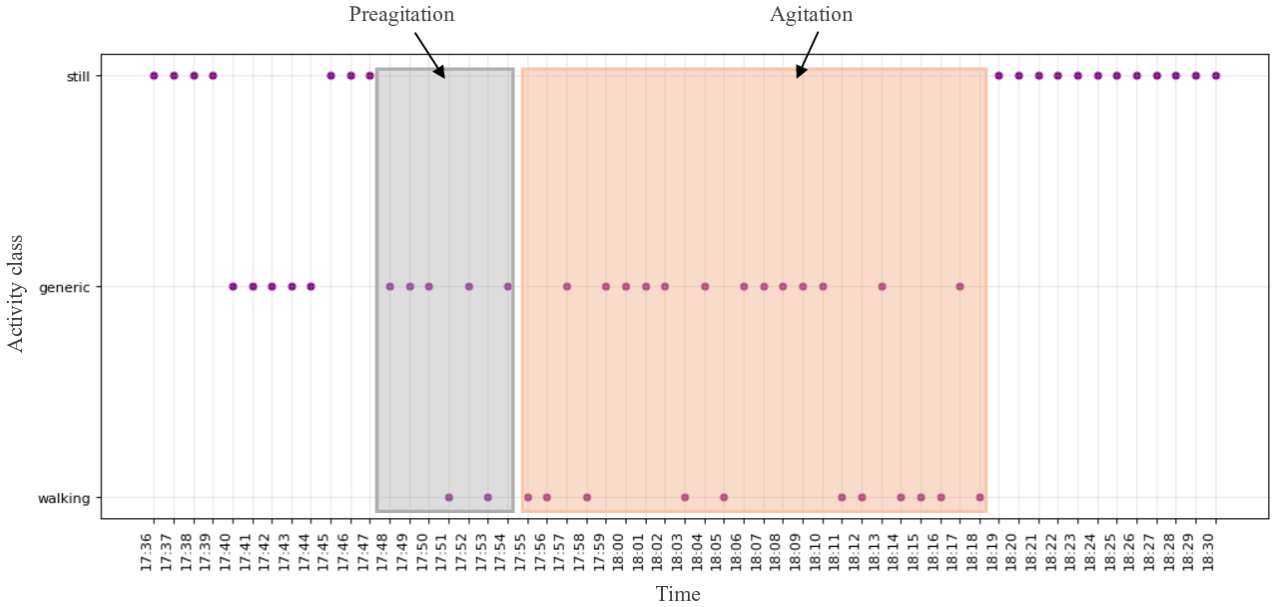


Figure 1 illustrates the activity class for participant #1, extracted from the accelerometer signal, revealing that the participant was in motion rather than stationary during AA and preagitation episodes, indicating body movement during these events. Figure 6 displays the total activity counts for participant #2 from the accelerometer signal. While the normal activity count for the participant ranged between 0 - 100 during AA and preagitation events, it surged to 50 - 140, signifying heightened activity levels during AA. Finally, Figure 7 presents the pulse rate derived from the heart rate signal for participant #3. Although the participant's average pulse rate ranged from 55 - 80 (SD 12.4) bpm, it increased to 90 - 110 (SD 11.6) bpm during AA and preagitation events. Across the raw and digital biomarkers data, we observed that the preagitation occurred from 3:20 PM to 3:27 PM. This indicates that signs of AA behavior occurred approximately 7 minutes before the actual event, suggesting the potential to predict and prevent AA events.

Performance of Video-Based Detection

We preprocessed our videos using OpenPose and performed feature extraction as described in the methodology section. As AA behaviors are repetitive in nature, we selected RNN models to capture the sequential patterns of these events. Features were extracted from 30-second windows. The window moves 1 second at a time to capture different variations of AA behaviors from the skeletal points. For the classification task, we tested 2 different network structures, which are LSTM and GRU.

In the feature analysis step, we focused on reducing the model's dimensionality without compromising its performance. Initially, 47 features were extracted based on skeletal movements,

including Euclidean distances and angles between key body parts. It is evident from the figures that more features are highly correlated in agitation events than in nonagitation events. This had a strong effect on the feature reduction as only the highly correlated features in both datasets were removed. We investigated the correlation of the features more deeply and tested the model on fewer features based on a correlation threshold. Setting the correlation threshold to 0.8 reduced the number of features to 39 features. The number of features based on the correlation threshold did not change even when the threshold was set to as low as 0.5 due to the huge difference in correlation between both types of events. Both datasets are randomly split into 70% for training and 30% for testing. The training was conducted on a Lambda server equipped with an RTX A6000 GPU, and each model was trained for 100 epochs with a batch size of 256. We used the Adam optimizer to efficiently handle sparse gradients and used a sigmoid activation function for binary classification (AA vs nonagitation). We report the results of all the tests on the testing set in Table 6.

Table . Video results comparison.

Model	Number of features	Accuracy	AUC[a]	F_1-score	Recall	Time (s)
LSTM[b]	47	0.94	0.98	0.97	0.96	16.1
LSTM	39	0.94	0.98	0.98	0.97	**14.7**[c]
GRU[d]	47	0.95[e]	0.98[e]	0.98[e]	0.97	30.2
GRU	39	0.95[e]	0.99[e]	0.98[e]	0.98[e]	29.8

[a]AUC: area under the curve.

[b]LSTM: long short-term memory.

[c]Bold text indicates the shortest inference time.

[d]GRU: gated recurrent unit.

[e]Bold text indicates the best result for each metric.

We compared metrics such as accuracy, AUC, and F_1-score. We also compared the response time, an essential factor in real-time applications, of all models. As observed in the table, the reduction in the number of features did not affect the performance of the models. The LSTM model, in both cases, achieved an accuracy of 94% and an AUC of 98%. The GRU model reached 95% accuracy and 99% AUC. Although the performance of both models is similar, the response time of the GRU model is double that of the LSTM. The response time for the LSTM was 15.9 seconds when all the features were used in training and was almost one second faster with fewer features. The GRU models, however, had a response time of 30.1 seconds for 47 features and 29.7 seconds for 39 features. The results show that the GRU model is superior across all evaluation metrics, albeit for the response time, where it lags behind the LSTM model by a huge margin. As we aim to detect AA as early as possible, the swifter response time can allow for timely interventions by health care providers in case of any AA event. The high AUC values of both models signify a strong ability to minimize the rate of false positives. This is crucial to ensure the reliability of the model in detecting real AA with lower false alarms, causing less overhead for health care providers. The faster response time posits the LSTM model as possibly the more advantageous model for real-time deployment.

Discussion

Principal Findings

The successful implementation of the system within the hospital setting, considering privacy and the positive feedback from patients and health care professionals, highlights the system's viability in a real-world clinical environment. The system used in this study integrated physiological data from the EmbracePlus wristband and video footage from CCTV cameras, allowing for a comprehensive and multimodal approach to AA detection. The EmbracePlus wristband system demonstrates promising results in detecting and classifying AA and preagitation events in individuals with severe dementia. The AA detection results are reflected in the video detection system, and the preagitation labels can be added to the system from EmbracePlus. The following discussion highlights key findings and their implications, followed by suggestions for future work.

The EmbracePlus wristband, which leverages both raw data and digital biomarkers, demonstrated its efficacy in discerning patterns associated with AA and preagitation. The personalized Extra Trees model emerged as the top-performing algorithm for the raw data, achieving high performance. Furthermore, features such as electrodermal tonic mean, accelerometer activity class, and pulse rate highlighted the significance of identifying AA. Furthermore, the outcomes of our analysis are promising and demonstrate the potential of predicting AA in dementia care settings in real time. We were able to predict preagitation events from all participants at least 6 minutes before the actual AA event. The identification of preagitation patterns in the data suggests that physiological changes precede observable AA behavior. Being the first to explore these patterns in individuals with severe dementia from the EmbracePlus wristband, this study lays the groundwork for a deeper understanding of the dynamics and physiological signatures of AA behaviors. The newfound ability to identify preagitation patterns offers a potential window for early intervention and preventive measures.

Moreover, the video detection system, incorporating CCTV footage and advanced pose estimation techniques, was used along with the EmbracePlus wristband for AA detection. The privacy preservation technique, which follows the REB

protocols in the hospital, does not exploit the patient's personal features or body image without affecting the performance of the proposed model. The LSTM neural network and the GRU networks exhibited robust performances. The high AUC of both models is particularly crucial in the context of health care to minimize the risk of false negatives and ensure that true AA events are accurately identified. During the real-time deployment stage, the model adapts and continuously improves based on the collected data. The labeled preagitation labels collected from the wristband can be fed into the training set of the video detection system to provide insight into detecting preagitation from the video footage. The outcomes of our analysis are promising and demonstrate the potential of both LSTM and GRU neural networks in detecting AA in dementia care settings in real time.

Future Work

For future work, we will focus on expanding our dataset by recruiting more participants from the hospital. We aim to validate the system over the long term, assessing its stability, generalizability, and adaptability to health care and home care environments. This step is crucial for building a database with a substantial number of AA and preagitation events, which is essential for developing a high-performance detection system using machine learning. Once our detection system is established, we plan to automate the real-time system that is capable of predicting AA. This involves receiving data in real time, classifying the data, and sending notifications to the health care providers if an AA event occurs. Additionally, we plan to improve the real-time AA detection of the video system. We also plan to use the preagitation labels from the EmbracePlus wristband to help our video detection model predict AA before they happen. During the initial stages of this study, health care providers will review and confirm all collected AA events. Their feedback is crucial in refining and enhancing the model's performance and should aid in identifying any limitations or challenges. Once the model is reliable, it will automatically detect and predict AA with no human intervention.

Conclusions

This study represents a notable step forward in developing an AA and preagitation detection system for individuals with severe dementia. It uses a comprehensive approach by integrating psychological biomarker sensing and video detection systems. The results demonstrate the feasibility and efficacy of monitoring systems that combine various data sources for AA detection. This study recruited 10 participants from the Ontario Shores Center for Mental Health Sciences Institute. We used the EmbracePlus wristband for continuous health monitoring and video footage from CCTV cameras for real-time observation of AA events. In the preliminary data analysis, the features extracted from the raw data of the EmbracePlus wristband demonstrated exceptional performance in detecting AA events, with the Extra Trees model emerging as the top-performing algorithm for all the personalized models and MLP outperforming the rest of the models for the general model and achieving an accuracy of 98%. Exploring the digital biomarkers further strengthened the system's classification of AA, preagitation, and normal events. Pulse rate, activity class, and activity counts have emerged as critical indicators for detecting AA. This study revealed the potential for detecting preagitation patterns, showcasing a 6-minute lead time before actual AA events. This early detection capability holds promise for timely intervention and preventive measures. In addition to the EmbracePlus wristband, the video-based detection demonstrated promising results in detecting AA using GRU, achieving a 95% accuracy rate and a robust AUC of 98%. The data analysis' promising results highlight the potential of the multimodal approach to enhance patient care and safety by predicting AA events. This research will provide new directions for researchers interested in technologies for dementia care and provide challenging propositions in detecting and monitoring, modeling, and evaluating patient-specific interventions for people living with dementia demonstrating NPS.

Acknowledgments

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Workflow of the data collection process.
[DOCX File, 280 KB - aging_v8i1e68156_app1.docx]

Multimedia Appendix 2
Feature importance for participants #1, #2, and #3.
[DOCX File, 104 KB - aging_v8i1e68156_app2.docx]

References

1. Dementia. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-06-10]

2. Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S. Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. JAMA 2002 Sep 25;288(12):1475-1483. [doi: 10.1001/jama.288.12.1475] [Medline: 12243634]
3. Ballard C, Corbett A. Agitation and aggression in people with Alzheimer's disease. Curr Opin Psychiatry 2013 May;26(3):252-259. [doi: 10.1097/YCO.0b013e32835f414b] [Medline: 23528917]
4. Seitz D, Purandare N, Conn D. Prevalence of psychiatric disorders among older adults in long-term care homes: a systematic review. Int Psychogeriatr 2010 Nov;22(7):1025-1039. [doi: 10.1017/S1041610210000608] [Medline: 20522279]
5. Knuff A, Leung RH, Seitz DP, Pallaveshi L, Burhan AM. Use of actigraphy to measure symptoms of agitation in dementia. Am J Geriatr Psychiatry 2019 Aug;27(8):865-869. [doi: 10.1016/j.jagp.2019.02.013] [Medline: 30952608]
6. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology (ECronicon) 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
7. Cohen-Mansfield J. Cohen-mansfield agitation inventory. Int J Geriatr Psychiatry 1986.
8. Artificial Intelligence in Healthcare Market Size & Trends Analysis Report. Research and Markets. 2022. URL: https://www.researchandmarkets.com/report/healthcare-ai?srsltid=AfmBOopaBJv8BVQKRVRLr1YHW1TDHwxLSyce4o5UHr0dB0DX9g43dh3N [accessed 2025-6-19]
9. Bharucha AJ, Anand V, Forlizzi J, et al. Intelligent assistive technology applications to dementia care: current capabilities, limitations, and future challenges. Am J Geriatr Psychiatry 2009 Feb;17(2):88-104. [doi: 10.1097/JGP.0b013e318187dde5] [Medline: 18849532]
10. Khan SS, Ye B, Taati B, Mihailidis A. Detecting agitation and aggression in people with dementia using sensors—a systematic review. Alzheimer's Dementia 2018 Jun;14(6):824-832. [doi: 10.1016/j.jalz.2018.02.004]
11. Fabrizio C, Termine A, Caltagirone C, Sancesario G. Artificial intelligence for Alzheimer's disease: promise or challenge? Diagnostics (Basel) 2021 Aug 14;11(8):1473. [doi: 10.3390/diagnostics11081473] [Medline: 34441407]
12. Goerss D, Köhler S, Rong E, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 7;7:e50107. [doi: 10.2196/50107] [Medline: 38848116]
13. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021 Dec 6;4(4):e30353. [doi: 10.2196/30353] [Medline: 34874886]
14. Ontario Shores Centre for Mental Health Sciences. 2022. URL: https://www.ontarioshores.ca [accessed 2025-06-10]
15. EmbracePlus. URL: https://www.empatica.com [accessed 2025-06-10]
16. Carrarini C, Russo M, Dono F, et al. Agitation and dementia: prevention and treatment strategies in acute and chronic conditions. Front Neurol 2021;12:644317. [doi: 10.3389/fneur.2021.644317] [Medline: 33935943]
17. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. a systematic review. Front Pharmacol 2019;10:1699. [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
18. Yang P, Bi G, Qi J, Wang X, Yang Y, Xu L. Multimodal wearable intelligence for dementia care in healthcare 4.0: a survey. Inf Syst Front 2025 Feb;27(1):197-214. [doi: 10.1007/s10796-021-10163-3]
19. Valembois L, Oasi C, Pariel S, Jarzebowski W, Lafuente-Lafuente C, Belmin J. Wrist actigraphy: a simple way to record motor activity in elderly patients with dementia and apathy or aberrant motor behavior. J Nutr Health Aging 2015 Aug;19(7):759-764. [doi: 10.1007/s12603-015-0530-z] [Medline: 26193860]
20. Etcher L, Whall A, Kumar R, Devanand D, Yeragani V. Nonlinear indices of circadian changes in individuals with dementia and aggression. Psychiatry Res 2012 Aug 30;199(1):77-78. [doi: 10.1016/j.psychres.2012.03.042] [Medline: 22521234]
21. Badawi A, Elgazzar K, Ye B, et al. Investigating multimodal sensor features importance to detect agitation in people with dementia. : IEEE; 2023 Presented at: 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE); Sep 24-27, 2023; Regina, SK, Canada p. 77-82. [doi: 10.1109/CCECE58730.2023.10288764]
22. Cheung JCW, So BPH, Ho KHM, Wong DWC, Lam AHF, Cheung DSK. Wrist accelerometry for monitoring dementia agitation behaviour in clinical settings: a scoping review. Front Psychiatry 2022;13:913213. [doi: 10.3389/fpsyt.2022.913213] [Medline: 36186887]
23. Chikhaoui B, Ye B, Mihailidis A. Ensemble learning-based algorithms for aggressive and agitated behavior recognition. : Springer Presented at: 10th International Conference, UCAmI 2016, Part II; Nov 29 to Dec 2, 2016; Canary Islands, Spain p. 9-20. [doi: 10.1007/978-3-319-48799-1_2]
24. Mishra PK, Iaboni A, Ye B, Newman K, Mihailidis A, Khan SS. Privacy-protecting behaviours of risk detection in people with dementia using videos. Biomed Eng Online 2023 Jan 21;22(1):4. [doi: 10.1186/s12938-023-01065-3] [Medline: 36681841]
25. Khan SS, Mishra PK, Javed N, et al. Unsupervised deep learning to detect agitation from videos in people with dementia. IEEE Access 2022;10:10349-10358. [doi: 10.1109/ACCESS.2022.3143990] [Medline: 35582495]
26. Marshall F. Video-based automatic detection of agitated behaviours Ph.D dissertation. : Ulster University; 2022.
27. Rose K, Specht J, Forch W. Correlates among nocturnal agitation, sleep, and urinary incontinence in dementia. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):78-84. [doi: 10.1177/1533317514524814] [Medline: 24670931]

28. Gong J, Rose KM, Emi IA, et al. Home wireless sensing system for monitoring nighttime agitation and incontinence in patients with Alzheimer's disease. : ACM; 2015 Oct 14 Presented at: Proceedings of the Conference on Wireless Health, ser WH '15; Oct 14-16, 2015; Bethesda, MD p. 5. [doi: 10.1145/2811780.2822324]
29. Khan SS, Zhu T, Ye B, et al. DAAD: a framework for detecting agitation and aggression in people living with dementia using a novel multi-modal sensor network. : IEEE; 2017 Presented at: 2017 IEEE International Conference on Data Mining Workshops (ICDMW); Nov 18-21, 2017; New Orleans, LA p. 703-710. [doi: 10.1109/ICDMW.2017.98]
30. Ye B, Khan SS, Chikhaoui B, et al. Challenges in collecting big data in a clinical environment with vulnerable population: lessons learned from a study using a multi-modal sensors platform. Sci Eng Ethics 2019 Oct;25(5):1447-1466. [doi: 10.1007/s11948-018-0072-y] [Medline: 30357559]
31. Zhang K, Yuan H, Fang Y. Evaluation of smart agitation prediction and management for dementia care and novel universal village oriented solution for integration, resilience, inclusiveness and sustainability. : IEEE; 2022 Presented at: 2022 6th International Conference on Universal Village (UV); Oct 22-25, 2022; Boston, MA p. 1-34. [doi: 10.1109/UV56588.2022.10185497]
32. Alam R, Gong J, Hanson M, et al. Motion biomarkers for early detection of dementia-related agitation. 2017 Jun 23 Presented at: MobiSys'17; Jun 23, 2017; Niagara Falls, NY p. 15-20 URL: https://dl.acm.org/doi/proceedings/10.1145/3089341 [doi: 10.1145/3089341.3089344]
33. AP Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Publishing, Inc; 2013.
34. Cummings J, Mintzer J, Brodaty H, et al. Agitation in cognitive disorders: International Psychogeriatric Association provisional consensus clinical and research definition. Int Psychogeriatr 2015 Jan;27(1):7-17. [doi: 10.1017/S1041610214001963] [Medline: 25311499]
35. Sclan SG, Reisberg B. Functional Assessment Staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4 Suppl 1(3):55-69. [doi: 10.1017/s1041610292001157] [Medline: 1504288]
36. AXIS M3077-PLVE network camera. Axis. 2024. URL: https://www.axis.com/products/axis-m3077-plve#compatible-discontinued [accessed 2025-06-10]
37. AXIS P3225-VE Mk II Network Camera - product support. Axis Communications. 2024. URL: https://www.axis.com/products/axis-p3225-ve-mk-ii/support [accessed 2025-06-10]
38. Badawi A, Choudhury S, Elgazzar K, Burhan AM. Artificial intelligence and features investigating to detect neuropsychiatric symptoms in patients with dementia: a pilot study. 2023 Presented at: 2023 IEEE Symposium Series on Computational Intelligence (SSCI); Dec 5-8, 2023; Mexico City, Mexico. [doi: 10.1109/SSCI52147.2023.10371835]
39. Cao Z, Martinez GH, Simon T, Wei S, Sheikh YA. OpenPose: Realtime multi- person 2d pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell 2019;43(1):172-186. [doi: 10.1109/TPAMI.2019.2929257]
40. Simon T, Joo H, Matthews I, Sheikh Y. Hand keypoint detection in single images using multiview bootstrapping. 2017 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI. [doi: 10.1109/CVPR.2017.494]
41. Cao Z, Simon T, Wei SE, Sheikh Y. Realtime multi-person 2d pose estimation using part affinity fields. 2017 Presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jul 21-26, 2017; Honolulu, HI. [doi: 10.1109/CVPR.2017.143]
42. Wei SE, Ramakrishna V, Kanade T, Sheikh Y. Convolutional pose machines. 2016 Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 27-30, 2016; Las Vegas, NV. [doi: 10.1109/CVPR.2016.511]

Abbreviations

AA: agitation and aggression
AI: artificial intelligence
AUC: area under the curve
GDU: Geriatric Dementia Unit
GRU: gated recurrent unit
JREB: Joint Research Ethics Board
LSTM: long short-term memory
MLP: multilayer perceptron
NPS: neuropsychiatric symptom
REB: research ethics board
RNN: recurrent neural network

Edited by R Yang; submitted 30.10.24; peer-reviewed by C Wu, S Kumar; revised version received 31.03.25; accepted 28.04.25; published 15.07.25.

Please cite as:
Badawi A, Elmoghazy S, Choudhury S, Elgazzar S, Elgazzar K, Burhan AM
Multimodal Detection of Agitation in People With Dementia in Clinical Settings: Observational Pilot Study
JMIR Aging 2025;8:e68156
URL: https://aging.jmir.org/2025/1/e68156
doi:10.2196/68156

Usability and Impact of the Web-Based Dementia Foundations Educational Program in Personal Support Workers (PSWs), PSW Trainees, and Care Companions: Quasi-Experimental Study

Anthony J Levinson[1*], MA, MSc, MD; Stephanie Ayers[1*], BA, BEd; Sandra Clark[1], MSc; Patricia Gerantonis[1], BSc, MPH; Amy Schneeberg[2], PhD; Richard Sztramko[3], BSc, MD

[1]Division of e-Learning Innovation, Faculty of Health Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

[2]Freelance Statistical Consultant, Vancouver, BC, Canada

[3]Department of Medicine, McMaster University, Hamilton, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Anthony J Levinson, MA, MSc, MD
Division of e-Learning Innovation, Faculty of Health Sciences, McMaster University, 1280 Main Street West, Hamilton, ON, Canada

Abstract

Background: Personal support workers (PSWs) are often expected to provide ongoing support for complex conditions and have identified an increased need for training in several areas, including dementia and mental health. Web-based interventions may be helpful complements to traditional in-person continuing education and training, but their effectiveness must be explored further.

Objective: This study's objective was to evaluate the usability, usefulness, satisfaction with, and effectiveness of the web-based Dementia Foundations Program among unregulated care providers who provide care to persons living with dementia or are in training.

Methods: A cohort of 50 PSWs, PSW trainees, and paid care companions from 3 recruitment sites were invited to access the Dementia Foundations Program, a 4-hour self-paced web-based program composed of 4 courses, for up to 6 weeks. Usability, usefulness, and satisfaction were assessed using surveys after each course and following the program. Dementia knowledge and attitudes were measured using the Dementia Knowledge Assessment Scale and the Dementia Attitudes Scale, with differences between baseline and postprogram scores analyzed using repeated measures ANOVA.

Results: Participants reported high levels of satisfaction with the program. Of the 50 participants, 46 (92%) agreed that the web-based training met their expectations, 47 (94%) agreed that the training covered a broad range of topics and was not missing any important content, and 49 (98%) agreed that the web-based training would benefit them. There was a significant postprogram improvement in dementia knowledge as measured by the Dementia Knowledge Assessment Scale, with an average 30% improvement across all cohorts. Dementia Attitudes Scale scores were also significantly improved postprogram across all cohorts.

Conclusions: This pilot study in PSWs, PSW trainees, and unregulated care companions demonstrated high satisfaction levels with the web-based Dementia Foundations Program. There were substantial improvements in knowledge and small improvements in attitudes for participants, and it was perceived as a useful tool that complemented their existing education and training. The Dementia Foundations Program is a user-friendly and effective e-learning program, which can be conveniently scaled and spread to enhance unregulated care provider dementia education.

(JMIR Aging 2025;8:e67889) doi:10.2196/67889

KEYWORDS

dementia; caregivers; e-learning; education; training; mobile phone

Introduction

Background

The number of older adults aged 60+ years is expected to double worldwide by 2050 [1]. In Canada specifically, the older seniors population aged 75+ years is expected to grow even sooner, possibly tripling by 2046 [2]. This has the potential to increase the prevalence of age-related health conditions such as dementia, as well as the need for adequate human health resources [3].

The Need for Unregulated Care Providers

There are estimated to be about 600,000 people in Canada living with dementia currently, with that number projected to grow to

almost a million in 2030 [4]. Unregulated care professionals who provide care to older adults play an important role in the effort to meet the growing demand for care from Canada's aging population. Several occupational titles can be used to describe these unregulated care providers, including personal support workers (PSWs), health care aides, and home care workers [5]. For the purposes of this paper, we will predominantly use the term PSWs, with the intent that this information applies to most professionals working in similar unregulated care roles.

It is estimated that there are over 1 million PSWs in Canada [6] providing support, including assisting with daily living activities, light housekeeping, meal preparation, socialization, and companionship [7], in several settings, including home, community, and clinical. In Canada, PSWs are not currently required to hold accredited qualifications and are not formally regulated by an organization or governing body. This lack of regulation of training can lead to gaps in knowledge and caregiving abilities [6]. PSWs have indicated that more training is needed, specifically in the field of dementia and mental health [8]. With over 69% of residents in long-term care (LTC) homes living with dementia and 87% with some form of cognitive impairment [9], more specific education and training about dementia are needed for PSWs.

Dementia Education Approaches

There are currently a variety of training approaches for dementia for unregulated care providers. Some PSW training programs or colleges might offer traditional didactic instruction on topics related to dementia; some of these programs have transitioned to webinar-based instruction during the pandemic. For PSWs or other unregulated care providers already in the workplace, the most common approach would be occasional on-the-job training, such as periodic "in-service" or "lunch-and-learn" sessions. These approaches tend to be synchronous, nonstandardized, and not necessarily educationally designed. Very few offer foundational knowledge content about dementia. Some approaches may also involve face-to-face instruction, which did prove to be challenging during the pandemic, concerning public health guidance as well as participant hesitance, even with the loosening of public health guidance [10]. Many of those solutions would be both challenging to scale and likely inconvenient for these care providers, who may not have flexible schedules that can accommodate synchronous instruction, especially if expected to be done on unpaid (personal) time.

For these reasons, the Dementia Foundations Program was developed in 2021 at McMaster University, aligning with several performance objectives and standards of the PSW Program Standard for Ontario Colleges from the Ministry of Training, Colleges and Universities [11]. During program development, project team members and target audience stakeholders identified the need for flexible, asynchronous solutions to best facilitate delivery to their PSW employees and students.

Program Description

The Dementia Foundations Program is a web-based, asynchronous (self-paced) series of 4 courses that takes approximately 4 hours total to complete. Each of the 4 courses includes modules on various topics related to dementia and dementia care and is described in Table 1.

Table . Dementia foundations program course descriptions.

Course number and course name	Course content
1. Foundations of Dementia	• What cognition is and how it is impacted by dementia • Difference between normal aging and mild cognitive impairment • Different types and stages of dementia • Types of treatments available
2. Responsive Behaviors and Mental Health	• Psychiatric issues such as apathy, depression, and anxiety that may affect people with dementia • Behavioral, emotional, and psychiatric symptoms of dementia • Strategies to help manage responsive behaviors
3. Home Supports and Safety	• Identifying safety risks • Strategies to reduce harm • Home and community-based supports and services available for patients and families
4. Promoting Brain Health and Caregiver Wellness	• Importance of brain health • How to make positive brain health choices • Why caregiver wellness is important and what can be done to support it

The program is intended as a complementary dementia education resource that may benefit care providers' training for several reasons. First, the modules within the courses use high-quality, evidence-based, multimedia instructional design [12,13], which has been shown to improve learning outcomes. Second, it can be delivered asynchronously for self-paced learning at any time and from any place, unlike other training programs that may use real-time training formats such as instructor-led classroom or webinar training. Third, the program has been developed using responsive web design and is compatible with any internet-connected device (smartphone, tablet, or desktop). Fourth, it is more efficient and cost-effective to scale than standardized digital or in-person training, as it does not require the presence of instructors. Lastly, the Dementia Foundations

Program offers a certificate of completion for each course and the complete program that learners can share with future employers detailing the time they have invested in learning about caring for those living with dementia.

Study Purpose

The objective of this study was to evaluate the usability, usefulness, satisfaction, and effectiveness of the Dementia Foundations Program among unregulated care providers who currently provide care to, or could provide care to, persons living with dementia.

Methods

Study Design

A pilot pre-post design was used with measures taken at baseline, after each of the 4 courses, and after program completion. Participants were able to complete the program at their own pace over a maximum of 6 weeks.

Participants and Recruitment Cohorts

A total of 50 participants were recruited through collaborative partner networks via an email campaign, keeping in line with typical pilot study sample size guidance in the literature [14,15]. Participants were organized into 3 recruitment cohorts: those working in the city of Ottawa's LTC homes, those currently a part of Durham College's PSW training program, and those recruited from the uCarenet homecare digital marketplace. Participants who fulfilled the following inclusion criteria were eligible to participate: (1) had a good command of the English language, (2) were sufficiently computer literate to use the Dementia Foundations Program, and (3) were currently an unregulated health care provider or trainee in Ontario, Canada. Interested participants were directed to a private web-based sign-up link where they were provided further study information and could submit informed consent. Once informed consent was received, a member of the research team manually added participants to the Dementia Foundations Program within the learning management system. Participants were grouped according to their recruitment cohort to allow the research team to identify and analyze patterns and trends among these 3 different settings.

The Dementia Foundations Program

Participants had access to the Dementia Foundations Program from June 1 to July 9, 2021. The program could be accessed from any internet-enabled device at any time or location. In addition to the 4 courses, participants in this pilot study had access to a "Getting to Know You" course at the beginning (which included the baseline assessments), and a "Review and Final Quiz" course at the end, which included the postprogram assessments.

Program Usability, Usefulness, and Satisfaction

To evaluate perceived usability, usefulness, and satisfaction with the training, 2 questionnaires were developed based on the Information Assessment Method for all questionnaire—a content-validated questionnaire designed to collect feedback from health information consumers based on 4 domains: situational relevance, cognitive impact, information use, and health benefits [16]. The first was administered after each course (postcourse assessment), and the second was delivered upon program completion (postprogram assessment). All questions were required, and each survey could only be submitted once; participants were able to review their responses and were asked to complete any questions they had missed before submitting.

The postcourse assessment consisted of 7 questions in total: 4 questions that could be answered on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree" ("I thought the course content was very important to my professional/learning needs" [usefulness], "I understood the content in this course" [usability], "I was able to complete the course in a reasonable amount of time" [usability], and "I would recommend this course to a colleague" [usefulness and satisfaction]), 2 questions that could be answered with multiple-select answers ("What do you think about this course?" [usefulness] and "Which benefit(s) are you expecting after taking this course?" [usefulness]), and 1 open text question ("Any other comments about this course?" [usability, usefulness, or satisfaction]).

The postprogram survey consisted of 8 questions total: 4 questions that could be answered on a 5-point Likert scale ranging from "strongly agree" to "strongly disagree" ("This online training met my expectations" [satisfaction], "I felt that this training covered a broad range of topics and was not missing any important content or topics" [usefulness and satisfaction], "This online training will benefit me" [usefulness and satisfaction], and "I feel my certificate/credential related to this program will be very valuable to me with current or future employers" [usefulness and satisfaction], 1 question that could be answered with a 3-point Likert scale ranging from "too low" to "too high" (thoughts on proposed pricing structure), 1 question that could be answered with multiple-select options ("Which of the following courses would you consider paying for?"), and 2 open text questions ("How much would you be willing to pay for this program?", and "Is there anything else you would like to say about this program?" (usability, usefulness, or satisfaction).

Dementia Knowledge

To evaluate the effectiveness of the Dementia Foundations Program, change in knowledge and attitudes related to dementia was assessed. Knowledge was measured at baseline and after program completion with the Dementia Knowledge Assessment Scale (DKAS; reliability α=.85; ω_h=0.87; overall scale) [17]. The DKAS consists of 25 items on different aspects of dementia that could be answered with "true," "probably true," "false," "probably false," or "I don't know." The total scores achievable for this scale range from 0 to 50, with greater dementia-related knowledge reflected by a higher score.

Dementia Attitudes

Attitudes about dementia were measured at baseline and after program completion with the Dementia Attitudes Scale (DAS) [18]. The DAS consists of 20 items on a 7-point Likert scale that reflect the affective, behavioral, and cognitive components of the attitudes toward individuals with Alzheimer disease and

related dementias. The total scores achievable for this scale range from 20 to 140, with a more positive attitude reflected by a higher score.

Learning Management System Setup

All assessments were offered digitally within a learning management system. The baseline "Getting to Know You" course included a demographic questionnaire (age, sex, race or ethnicity, highest level of education completed, occupation, and most common workplace setting), the DKAS, and the DAS. The postcourse assessment was built directly into the end of each of the 4 courses. The postprogram "Review and Final Quiz" course included the DKAS, DAS, and postprogram assessment.

Data Analysis

Descriptive analyses were used to summarize the baseline characteristics of the study population. Differences between baseline and postintervention DKAS and DAS scores for all groups were analyzed using repeated measures ANOVA. All analyses were performed using R (version 4.4.1; R Foundation; June 14, 2024).

Ethical Considerations

The Hamilton Integrated Research Ethics Board reviewed the study protocol and granted exemption from full review on March 18, 2021, as this was considered a quality improvement initiative. Participants were required to provide informed consent. All participants were informed of the length of time of the e-learning and surveys, as well as details surrounding data collection, storage, and investigator identities. Participants' identities and confidentiality were maintained throughout the research study. All participant data were deidentified and were stored on password-protected secure servers to prevent unauthorized access. There was no known risk or harm to participating in this study or publicizing its results or findings. Upon program completion, participants were provided with a CAD $80 (US $58.22) gift card and a certificate of completion for their professional portfolio.

Results

Participant Characteristics

A total of 50 unregulated care providers enrolled and completed this study. Of these, 24 (48%) participants identified as PSWs, 16 (32%) identified as a PSW trainee or student, 1 (2%) identified as a health care aide, 1 (2%) identified as a personal care assistant, 4 (8%) identified as other front line health care workers, 1 (2%) identified as unregulated care provider manager, and 3 (6%) identified as other (while with no further details on what their role was, care settings identified were retirement home or LTC [recruited from the Durham College cohort], group home [recruited from the uCarenet cohort], and possibly for a family member [recruited from the uCarenet cohort] for these 3 participants). Participant demographics broken down by the 3 recruitment cohorts are outlined in Table 2.

Table . Participant demographics in each recruitment cohort.

Characteristics	Number of responses per recruitment cohort		
	City of Ottawa LTC[a] (n=20), n (%)	Durham College (n=17), n (%)	uCarenet (n=13), n (%)
Age (years)			
18 to 24	1 (5)	2 (12)	0 (0)
25 to 34	2 (10)	9 (53)	5 (38)
35 to 44	9 (45)	3 (18)	4 (31)
45 to 54	5 (25)	3 (18)	2 (15)
55 to 64	3 (15)	0 (0)	2 (15)
Highest level of education completed			
High school	0 (0)	2 (12)	0 (0)
College or university	17 (85)	15 (88)	13 (100)
Graduate school	3 (15)	0 (0)	0 (0)
Role			
PSW[b]	19 (95)	0 (0)	5 (38)
PSW trainee or student	0 (0)	16 (94)	0 (0)
Health care aide	1 (5)	0 (0)	0 (0)
Personal care assistant	0 (0)	0 (0)	1 (8)
Other front-line health care worker	0 (0)	0 (0)	4 (31)
Unregulated care provider manager or administrator	0 (0)	0 (0)	1 (8)
Other	0 (0)	1 (6)	2 (15)

[a]LTC: long-term care.

[b]PSW: personal support worker.

Program Usability, Usefulness, and Satisfaction

Overall, the courses were positively rated, with most participants agreeing that the course was “very important” to their professional and learning needs (Table 3). Participants agreed that the training benefited them by “improving the health and well-being of the people that they care for.” Additionally, participants identified that the training allowed them to “feel more confident,” “prevent a problem,” “handle a problem or the worsening of a problem,” and “decide something with someone else.” The participants were also asked to provide any additional comments about the course; many open-text responses were very positive. Participants identified the importance and applicability of the information to their professional work.

In general, participants reported high levels of satisfaction with the program. A total of 46 of 50 (92%) of participants agreed or strongly agreed that the web-based training met their expectations, 47 (94%) agreed or strongly agreed that the training covered a broad range of topics and was not missing any important content, 49 (98%) agreed or strongly agreed that the web-based training would benefit them, and 46 (92%) agreed or strongly agreed that the program would be very valuable with current or future employers. No support inquiries were received from participants, further confirming the high usability of the program.

Table . Summary of postcourse feedback.

Question and course	Response breakdown, n (%)				
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I thought the course content was very important to my profession/learning needs					
Course 1 (Foundations of Dementia)	4 (8)	0 (5)	1 (2)	18 (36)	27 (54)
Course 2 (Responsive Behaviours and Mental Health)	0 (0)	0 (0)	0 (0)	20 (40)	30 (60)
Course 3 (Home Supports and Safety)	1 (2)	1 (2)	1 (2)	19 (38)	28 (56)
Course 4 (Promoting Brain Health and Caregiver Wellness)	0 (0)	1 (2)	1 (2)	28 (56)	20 (40)
I was able to complete the course in a reasonable amount of time.					
Course 1 (Foundations of Dementia)	0 (0)	0 (0)	7 (14)	24 (48)	19 (38)
Course 2 (Responsive Behaviours and Mental Health)	0 (0)	0 (0)	2 (4)	29 (58)	19 (38)
Course 3 (Home Supports and Safety)	0 (0)	1 (2)	5 (10)	22 (44)	22 (44)
Course 4 (Promoting Brain Health and Caregiver Wellness)	1 (2)	0 (0)	3 (6)	28 (56)	18 (36)
I would recommend this course to a colleague.					
Course 1 (Foundations of Dementia)	0 (0)	0 (0)	2 (4)	13 (25)	35 (70)
Course 2 (Responsive Behaviours and Mental Health)	1 (2)	0 (0)	2 (4)	19 (38)	28 (56)
Course 3 (Home Supports and Safety)	0 (0)	0 (0)	0 (0)	19 (38)	31 (62)
Course 4 (Promoting Brain Health and Caregiver Wellness)	0 (0)	0 (0)	1 (2)	20 (40)	29 (58)

Change in Dementia Knowledge

All 50 participants completed the baseline and postprogram DKAS. One participant from the City of Ottawa LTC recruitment cohort was missing data for 15/25 items on the baseline DKAS due to a technical glitch. The most conservative approach was taken where it was assumed the participant would have received 2 points for each missing answer (ie, the highest score possible). The average baseline DKAS score was 34.2 of 50 (68%, SD 10.5) and the average postprogram score was 44.2 of 50 (88.4%, SD 4.7). Scores for all participants are shown in Figure 1, with scores for each recruitment cohort shown in Figure 2.

Figure 1. Baseline and postprogram DKAS scores for all participants. Results of repeated measures ANOVA (mean with 95% CI) with time as the only predictor of DKAS, including a random effect for participant ID. *Represents statistical significance (*P*<.001). DKAS: Dementia Knowledge Assessment Scale.

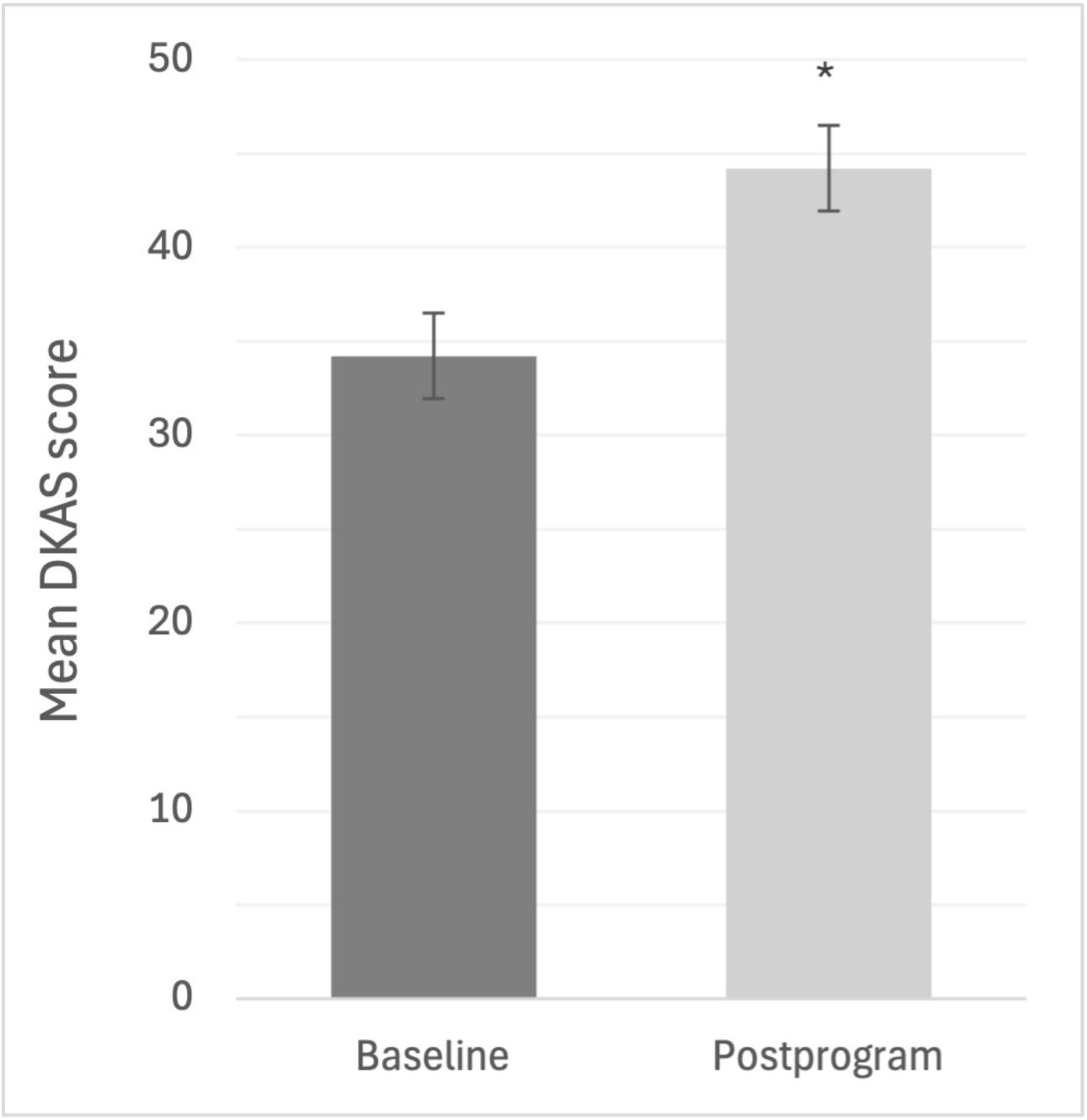
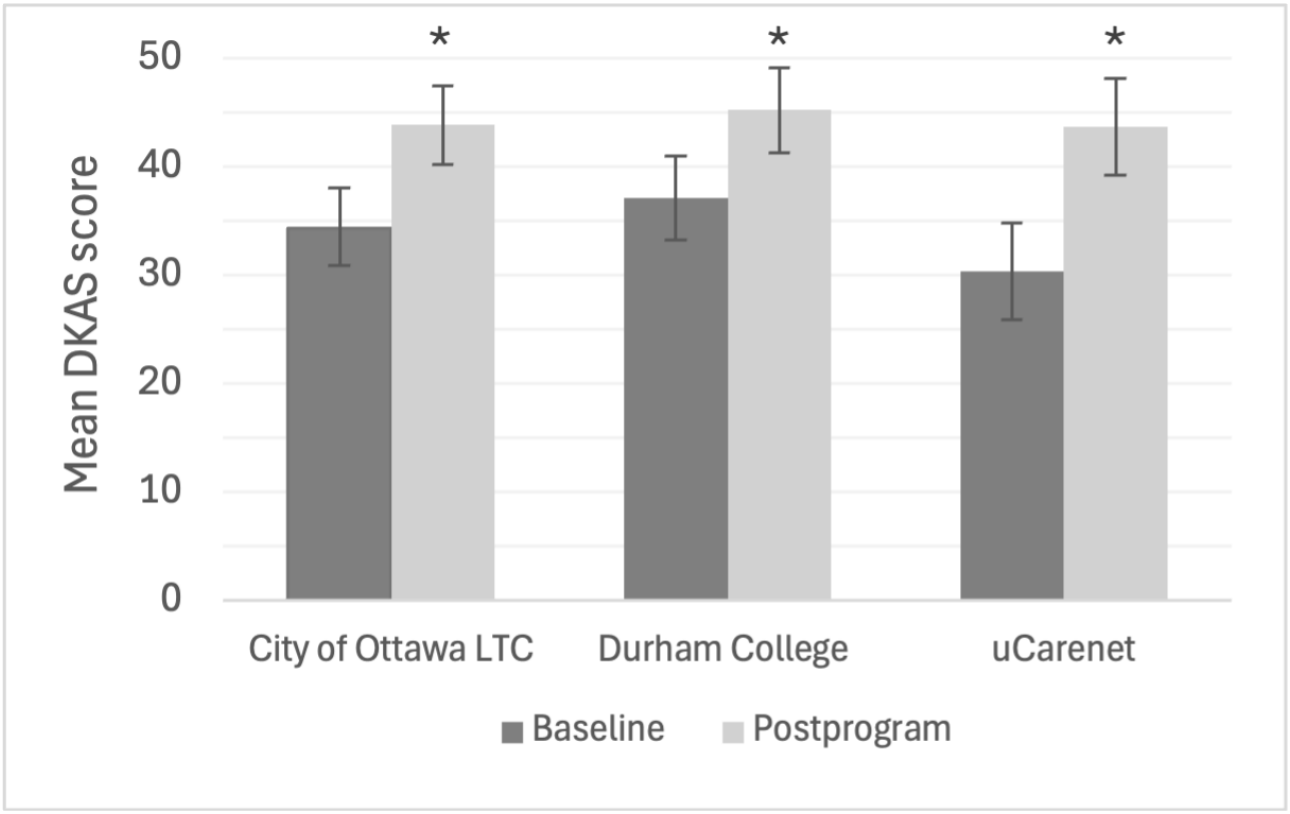


Figure 2. Baseline and postprogram DKAS by cohort and time. Results of repeated measures DKAS model (mean with 95% CI) with cohort time interaction, including a random effect for participant ID. *Represents statistical significance (P<.05). DKAS: Dementia Knowledge Assessment Scale; LTC: long-term care.

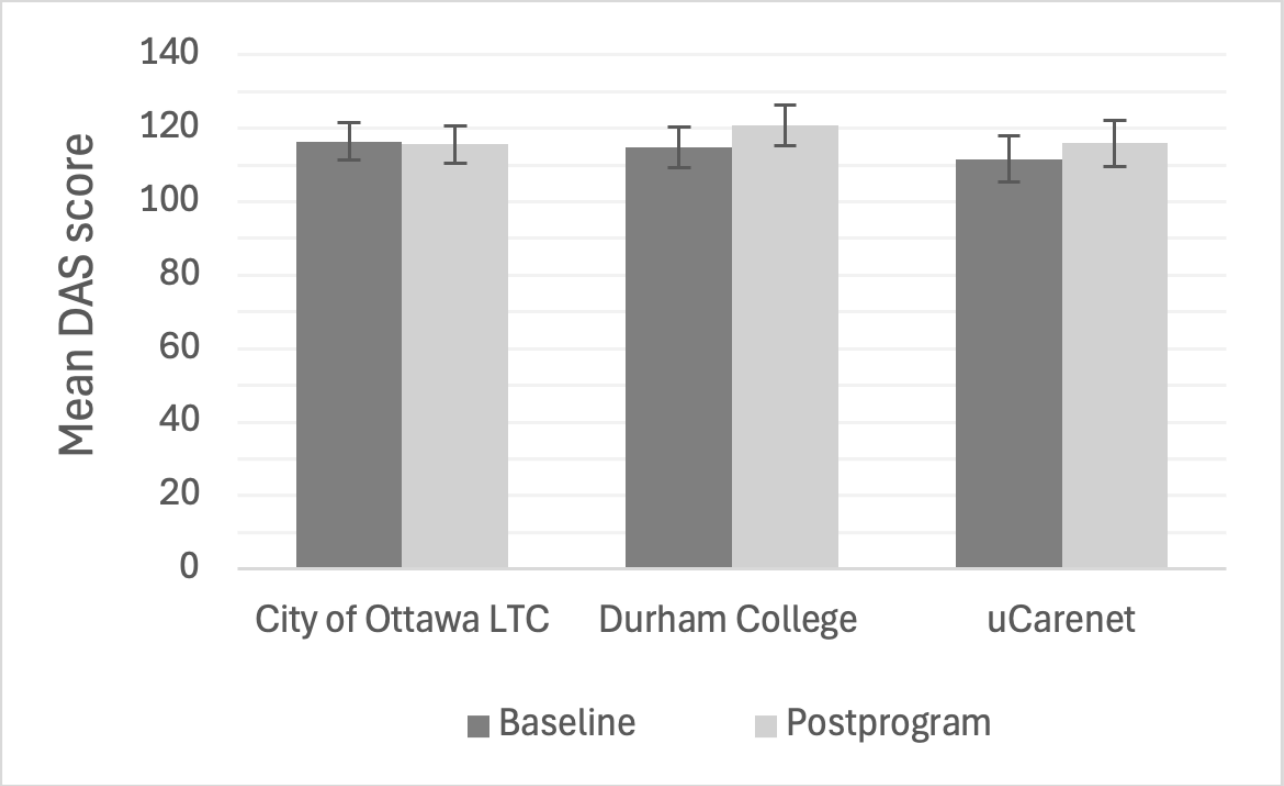


Repeated measures ANOVA showed a statistically significant effect on DKAS scores ($F_{1,49}$=58.533, P<.001), with scores increasing from baseline to postprogram across all 3 recruitment cohorts (η2=0.276). The model, including only time as a predictor and a random effect for participant ID, demonstrated that relative to baseline, there was an average 9.98-point increase on the DKAS score postintervention (95% CI 7.88-12.08; SE 1.065; P<.001). Post hoc testing did not reveal any significant differences in DKAS scores between recruitment cohorts; the model including an interaction with time and cohort, and a model with just time (not including cohort at all), were not statistically significantly different (P=.24).

Change in Dementia Attitudes

All 50 participants completed the baseline and postprogram DAS. On average across all recruitment cohorts, participants scored 114.7 of 140 (82%; SD 11.6) at baseline and 117.5 of 140 (84%; SD 11.2) postprogram. Scores for each recruitment cohort are shown in Figure 3.

Figure 3. Baseline and postprogram DAS scores among participants by cohort. Results from repeated measures DAS model (mean and 95% CI) including interaction for cohort and time. DAS: Dementia Attitudes Scale; LTC: long-term care.



Repeated measures ANOVA showed a statistically significant effect on DAS scores ($F_{1,49}$=4.383, P=.04; η2= 0.015). The model including only time as a predictor demonstrated that relative to baseline, there was an average 2.8 point increase on the DAS score postprogram (P=.04). Post hoc testing did not reveal any significant differences in DAS scores between recruitment cohorts; the model including an interaction with time and cohort, and a model with just time (not including cohort at all) were not statistically significantly different (P=.15).

Discussion

Principal Findings

The purpose of this pilot study was to evaluate the usability, usefulness, satisfaction, and effectiveness of the Dementia Foundations Program, a series of 4 web-based courses about various aspects of dementia developed for unregulated care providers such as PSWs. Three sites were recruited from slightly different populations (PSW trainees, PSWs working in LTC, and PSWs and other unregulated care providers working in community settings) who were all willing to take part in recruitment efforts due to the perceived benefit of additional dementia care training for their students, employees, and members. Overall, the program was positively valued for its usability, and participants reported a high level of satisfaction with the program. Participant knowledge, as measured by the DKAS, significantly increased after completing the program, with an average 30% improvement across all recruitment cohorts. A statistically significant increase was also seen for dementia attitudes as measured by the DAS, with a more modest average improvement of 2.5%. While scores for both the DKAS and DAS did not significantly differ between cohorts, there are a few trends worth highlighting.

PSW trainees from the Durham College cohort had higher baseline DKAS scores and slightly lower DAS scores compared to the PSWs from both the City of Ottawa LTC and uCarenet cohorts. The differences in knowledge scores could be explained by how trainees might have higher fact-based knowledge (from more recent studying) and may be more used to taking knowledge assessments, whereas practicing PSWs may have more everyday clinical knowledge and may not have taken a formal course in some time. When looking at dementia attitudes, the overall significant increase can be attributed to the Durham College and uCarenet cohorts, rather than the Ottawa LTC cohort. This conceptually makes sense, as practicing PSWs may already possess more developed attitudes toward dementia that would not be expected to change drastically after only 4 hours of knowledge-based education. Conversely, PSW trainees and PSWs not working in LTC homes who have had less exposure to working with people with dementia may be more prone to shifting their attitude in a significant way after 4 hours of education.

Comparison to Prior Work

Dementia Education

There is currently a lack of dementia-specific care training offered to unregulated care providers, leading to gaps in care [8,19]. This highlights the need for increased and enhanced training to improve education with the purpose of improving care for individuals with dementia. Research has shown that dementia education and training for health care workers improves knowledge, fosters positive attitudes, increases

confidence, and ultimately produces better outcomes for individuals living with dementia [20,21]. An extensive review by Surr et al [22] evaluated various features of dementia education programs for health care providers. Among the findings, the authors concluded that effective dementia education programs should be relevant to the workers' role and experience, include practice-based learning with theoretical content, be delivered by an experienced facilitator, and involve in-person facilitation. Although there are some existing on-the-job dementia training programs for unregulated care providers, they vary in availability, are more expensive to deliver, and are often less convenient for care providers who may not have flexible schedules to accommodate this type of instruction. The COVID-19 pandemic also imposed limitations on access to many in-person education programs. Thus, a blended knowledge approach that combines in-person training with web-based learning can deliver feasible and effective dementia education with the potential to be scaled and spread to a large population.

Furthermore, the recommendation by Surr et al [22] for in-person delivery may be influenced by several factors. For example, it is likely the case that most education and training programs about dementia have historically been delivered in person. Second, those programs delivered digitally (whether synchronously through webinar platforms or asynchronously) may not have incorporated best practices concerning the instructional design of multimedia e-learning. High-quality instructional design has been shown to improve learning transfer. The Dementia Foundations Program was created using evidence-based principles for multimedia learning outlined by Mayer [12] and Clark and Mayer [13]. The findings from the current study indicate that this entirely asynchronous web-based program is a useful and effective training modality for improving dementia-related knowledge among unregulated care providers. The high satisfaction with the program is evident in the course open-text comments, where most participants praised the program for its informativeness, user-friendliness, and usefulness to their profession as unregulated care providers.

About DKAS

Significant improvements in dementia knowledge were seen after the intervention for both PSWs and PSW trainees. As unregulated care providers in Canada have noted a lack of dementia-specific care training and education, this indicates that the Dementia Foundations Program is an effective way to impact dementia knowledge for this group. Other web-based dementia education programs have also been shown to be effective in increasing dementia knowledge. Eccleston et al [23] evaluated the efficacy of the 9-week-long Understanding Dementia Massive Open Online Course (UDMOOC) for people with varying professional and educational backgrounds. Similar to the Dementia Foundations Program, the UDMOOC aims to improve dementia knowledge by teaching a broad community audience basic neurobiology, dementia pathophysiology, medical management, and person-centered care. The median DKAS score at baseline for their study was 34.5 out of 50, which is comparable to our current pilot study's baseline median DKAS score of 35.5 out of 50, implying both groups started with similar levels of dementia knowledge before embarking on the training. The median postprogram scores for both studies were 45 of 50, both demonstrating a significant increase in dementia knowledge. These results are consistent with the current study's findings, indicating that web-based dementia education modalities are effective at increasing dementia knowledge. However, it is worth noting that the Dementia Foundations Program can be completed in a single day with only 4 hours' worth of content, whereas the UDMOOC estimates 21 hours over a 7- to 9-week period. Comparable baseline and postprogram median DKAS scores for both programs suggest the Dementia Foundations Program is more efficient for busy professionals, with similar knowledge gain in a much shorter amount of time.

When considering the mode of delivery of instruction, 1 study conducted by Parveen et al [24] also found that the most impactful dementia training programs involved a blend of both e-learning and in-person delivery, although another study conducted by Vollmar et al [25] found no significant differences between learning through in-person compared with web-based learning. With the significant increase in dementia knowledge shown by both the current study and Eccleston et al using only a web-based approach, it is reasonable to consider digital-only approaches to some aspects of dementia education, especially when considering the increased flexibility this offers to an unregulated care worker audience with potentially variable and less flexible schedules. A blended approach has been shown to be effective in other studies, and one can speculate that an efficient and effective approach might involve high-quality asynchronous web-based approaches for baseline knowledge education, complemented by in-person skills training.

About DAS

Dementia attitudes significantly increased from baseline to postprogram, with scores moving from an average of 114.7 to 117.5. While this increase is modest at 2.5%, baseline scores on the DAS started quite high compared to other studies looking at dementia attitudes in nursing students in Malta (mean DAS score of 107.9) [26] and health care professionals in China (mean DAS score of 91.3) [27]. The highest possible score on the DAS is 140, corresponding to the highest possible positive attitude toward dementia. The DAS scores from the current study indicate that this sample of participants already possesses relatively positive attitudes toward dementia compared to previous studies, therefore leaving only modest room for improvement on this measure.

Baseline DAS scores averaged 114.7 across all 3 recruitment cohorts, with uCarenet starting with the lowest score of 111.6 and the City of Ottawa LTC starting with the highest score at 116.4. After the 4-hour intervention, DAS scores were significantly higher at 117.5 across all 3 recruitment cohorts, with the Durham College PSW trainees showing the largest increase among cohorts from 114.9 to 120.9. Logically, this cohort showed the biggest increase in positive attitudes toward dementia, as trainees have very likely spent less time working with older adults and those with dementia. The experienced PSWs and other paid caregivers that comprised the Ottawa LTC and uCarenet cohorts have likely been working with people with dementia for many years and have had ample time to develop personal attitudes toward dementia, as previous research

has shown that working with and having ongoing and meaningful interactions with individuals living with dementia are key to fostering a change in attitudes [26,28]. It can therefore be reasonably assumed that 4 hours' worth of independent learning is not a sufficient intervention length to cause an attitude change in experienced workers, but remains sufficient exposure to improve attitudes toward dementia among trainees. This contrasts with a study that explored the impact of a single dementia awareness session on changing dementia attitudes among adolescents, with no significant improvements in either the intervention or control groups [29]. However, the difference may be that the Durham College PSW trainees were more accepting of a potential attitude change toward dementia given their desired career choice as a PSW.

Limitations

We acknowledge that there are limitations to this study. First, the relatively small sample size of 50 PSWs, PSW trainees, and other unregulated care professionals from Ontario who took part in this study is not necessarily representative of the diverse and large population of unregulated care providers in Canada or internationally. However, as the sample size was based on the typical large effect sizes for e-learning interventions, the results of this study remain valuable for an initial assessment of a new e-learning program for this population. This was further confirmed by the detection of statistically significant changes in both dementia knowledge and attitudes as measured by the DKAS and DAS. Second, the lack of a control group due to the pre-post study design introduces an unknown risk of bias; however, the pre-post design does alleviate intraparticipant variability and is consistent with the majority of studies in a 2019 systematic review of studies assessing technology-delivered dementia education to health care providers [30]. Lastly, not all participants had active experience working with people living with dementia, and as a result, may not have found the learning content to be relevant. However, trainees were intentionally included to be able to assess the program's suitability to incorporate into a training program, rather than continuing education once employed.

Conclusions and Future Directions

The Dementia Foundations Program was positively evaluated by a group of 50 unregulated care providers, predominantly PSWs and PSW trainees, with participants showing a significant improvement in both dementia knowledge and attitudes. Participants found the training very useful in supplementing their existing education and training, with particular relevance to the PSW trainee cohort. The findings of this study have implications for unregulated care providers who support people living with dementia, as well as trainees who may not yet interact with those living with dementia. The Dementia Foundations Program is an effective and user-friendly e-learning program that can be conveniently incorporated within existing dementia training and educational programs as part of a blended teaching and learning strategy to enhance dementia knowledge in unregulated care providers and trainees. Widespread implementation has the potential to increase training capacity and fill gaps in existing workforce training and improve dementia knowledge, attitudes, and awareness among care providers as dementia prevalence increases in Canada.

Since this pilot study was conducted, the Dementia Foundations Program has reached 1012 users from December 2021 to September 2024, with approximately 55% completing all 4 courses and receiving a certificate. Postprogram evaluations show high satisfaction with the courses from individuals who purchased the program digitally and employees and students from organizations who purchased the program for professional development.

Future work will focus on the spread and scale of this training program through advertising and partnerships with relevant organizations and institutions, and future research should use implementation science to assess organizational adoption and effectiveness. Continuous quality improvement evaluations will be conducted to determine the best approach for widespread implementation.

Acknowledgments

The authors disclose receipt of the following financial support for the research, authorship, or publication of this paper: This project received funding from the Centre for Aging & Brain Health Innovation (CABHI). AJL receives support through the John R Evans Chair in Health Sciences Educational Research and Instructional Development, Faculty of Health Sciences, McMaster University. The authors wish to thank Sheila Bauer, BScN, MHA, and the City of Ottawa LTC, Deborah Schuh, BScN, and the Durham College PSW Training Program, and Nectari Charitakis, BA, and uCarenet for their collaboration and assistance with recruitment. The authors also wish to thank Dr David Streiner for his guidance and mentorship throughout the statistical analysis.

Data Availability

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

Authors' Contributions

AJL and SA designed this study and led recruitment and data collection. SA, AJL, and AS performed data analyses. SA, SC, and PG drafted this paper. AS and SC created the figures. All authors read and approved this final paper.

Conflicts of Interest

AJL and RS are co-owners of the Dementia Foundations Program. There is a fee to access the Dementia Foundations Program.

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-06-17]
2. A portrait of canada's growing population aged 85 and older from the 2021 census. Statistics Canada. URL: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021004/98-200-X2021004-eng.cfm [accessed 2025-06-17]
3. Nuernberger K, Atkinson S, MacDonald G. Seniors in transition: exploring pathways across the care continuum. Healthc Q 2018 Apr;21(1):10-12. [doi: 10.12927/hcq.2018.25529] [Medline: 30051809]
4. Dementia numbers in Canada. Alzheimer Society. URL: https://alzheimer.ca/en/about-dementia/what-dementia/dementia-numbers-canada [accessed 2025-06-17]
5. PSW positions and job titles. Caring Support. URL: https://www.caringsupport.com/blog/consider-these-10-alternative-psw-positions-and-job-titles [accessed 2025-06-17]
6. Zagrodney K, Saks M. Personal support workers in Canada: the new precariat? Healthc Policy 2017 Nov;13(2):31-39. [doi: 10.12927/hcpol.2017.25324] [Medline: 29274225]
7. PSW roles & responsibilites. Ontario Personal Support Workers Association. URL: https://ontariopswassociation.com/psw-roles-and-responsibilities [accessed 2025-06-17]
8. Lum J, Sladek J, Ying A. Ontario personal support workers in home and community care: CRNCC/PSNO survey results. : CRNCC/RCRSC; 2009 URL: https://www.torontomu.ca/content/dam/crncc/knowledge/infocus/factsheets/InFocus-Ontario%20PSWs%20in%20Home%20and%20Community%20Care.pdf [accessed 2025-06-17]
9. Dementia in Canada: summary. Canadian Institute for Health Information. 2016. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-in-canada-summary [accessed 2025-07-03]
10. Hapsari AP, Ho JW, Meaney C, et al. The working conditions for personal support workers in the Greater Toronto Area during the COVID-19 pandemic: a mixed-methods study. Can J Public Health 2022 Dec;113(6):817-833. [doi: 10.17269/s41997-022-00643-7] [Medline: 35616873]
11. Personal support worker standard. : Ontario Ministry of Colleges and Universities; 2022 URL: https://files.ontario.ca/mcu-personal-support-worker-jan2022-41469-en-2022-03-31.pdf [accessed 2025-06-17]
12. Mayer RE. Multimedia learning. In: Psychology of learning and motivation: Elsevier; 2002, Vol. 41:85-139.
13. Clark RC, Mayer RE. E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning: John Wiley & Sons; 2016.
14. Lewis M, Bromley K, Sutton CJ, McCray G, Myers HL, Lancaster GA. Determining sample size for progression criteria for pragmatic pilot RCTs: the hypothesis test strikes back!. Pilot Feasibility Stud 2021 Feb 3;7(1):40. [doi: 10.1186/s40814-021-00770-x] [Medline: 33536076]
15. In J. Introduction of a pilot study. Korean J Anesthesiol 2017 Dec;70(6):601-605. [doi: 10.4097/kjae.2017.70.6.601] [Medline: 29225742]
16. Pluye P, El Sherif R, Granikov V, et al. Health outcomes of online consumer health information: a systematic mixed studies review with framework synthesis. J Assoc Inf Sci Technol 2019 Jul;70(7):643-659. [doi: 10.1002/asi.24178] [Medline: 31423458]
17. Annear MJ, Toye C, Elliott KEJ, McInerney F, Eccleston C, Robinson A. Dementia Knowledge Assessment Scale (DKAS): confirmatory factor analysis and comparative subscale scores among an international cohort. BMC Geriatr 2017 Jul 31;17(1):168. [doi: 10.1186/s12877-017-0552-y] [Medline: 28760154]
18. O'Connor ML, McFadden SH. Development and psychometric validation of the Dementia Attitudes Scale. Int J Alzheimers Dis 2010;2010:1-10. [doi: 10.4061/2010/454218]
19. Janes N, Sidani S, Cott C, Rappolt S. Figuring it out in the moment: a theory of unregulated care providers' knowledge utilization in dementia care settings. Worldviews Evid Based Nurs 2008;5(1):13-24. [doi: 10.1111/j.1741-6787.2008.00114.x] [Medline: 18333923]
20. Liu JYW, Lai C, Dai D, Ting S, Choi K. Attitudes in the management of patients with dementia: comparison in doctors with and without special training. East Asian Arch Psychiatry 2013 Mar;23(1):13-20. [Medline: 23535628]
21. Jennings A, McLoughlin K, Boyle S, et al. Development and evaluation of a primary care interprofessional education intervention to support people with dementia. J Interprof Care 2019;33(5):579-582. [doi: 10.1080/13561820.2018.1541876] [Medline: 30422731]
22. Surr CA, Gates C, Irving D, et al. Effective dementia education and training for the health and social care workforce: a systematic review of the literature. Rev Educ Res 2017 Oct;87(5):966-1002. [doi: 10.3102/0034654317723305] [Medline: 28989194]
23. Eccleston C, Doherty K, Bindoff A, Robinson A, Vickers J, McInerney F. Building dementia knowledge globally through the understanding dementia Massive Open Online Course (MOOC). NPJ Sci Learn 2019;4(1):3. [doi: 10.1038/s41539-019-0042-4] [Medline: 30993003]

24. Parveen S, Smith SJ, Sass C, et al. Impact of dementia education and training on health and social care staff knowledge, attitudes and confidence: a cross-sectional study. BMJ Open 2021 Jan 19;11(1):e039939. [doi: 10.1136/bmjopen-2020-039939] [Medline: 33468498]
25. Vollmar HC, Mayer H, Ostermann T, et al. Knowledge transfer for the management of dementia: a cluster randomised trial of blended learning in general practice. Implement Sci 2010 Jan 4;5(1):1. [doi: 10.1186/1748-5908-5-1] [Medline: 20047652]
26. Scerri A, Scerri C. Nursing students' knowledge and attitudes towards dementia - a questionnaire survey. Nurse Educ Today 2013 Sep;33(9):962-968. [doi: 10.1016/j.nedt.2012.11.001] [Medline: 23182308]
27. Zhao W, Moyle W, Wu MLW, Petsky H. Hospital healthcare professionals' knowledge of dementia and attitudes towards dementia care: a cross-sectional study. J Clin Nurs 2022 Jul;31(13-14):1786-1799. [doi: 10.1111/jocn.15590] [Medline: 33295010]
28. Tan WJ, Yeo D, Koh HJ, Wong SC, Lee T. Changing perceptions towards dementia: how does involvement in the arts alongside persons with dementia promote positive attitudes? Dementia (London) 2021 Jul;20(5):1729-1744. [doi: 10.1177/1471301220967805] [Medline: 33100055]
29. Farina N, Hughes LJ, Jones E, et al. The effect of a dementia awareness class on changing dementia attitudes in adolescents. BMC Geriatr 2020 Jun 2;20(1):188. [doi: 10.1186/s12877-020-01589-6] [Medline: 32487024]
30. Scerbe A, O'Connell ME, Astell A, Morgan D, Kosteniuk J, DesRoches A. Digital tools for delivery of dementia education for health-care providers: a systematic review. Educ Gerontol 2019 Nov 2;45(11):681-699. [doi: 10.1080/03601277.2019.1687149]

Abbreviations

DAS: Dementia Attitudes Scale
DKAS: Dementia Knowledge Assessment Scale
LTC: long-term care
PSW: personal support worker
UDMOOC: Understanding Dementia Massive Open Online Course

Edited by A Garcia; submitted 24.10.24; peer-reviewed by B Davis, M Sengupta; revised version received 22.05.25; accepted 26.05.25; published 21.07.25.

Please cite as:
Levinson AJ, Ayers S, Clark S, Gerantonis P, Schneeberg A, Sztramko R
Usability and Impact of the Web-Based Dementia Foundations Educational Program in Personal Support Workers (PSWs), PSW Trainees, and Care Companions: Quasi-Experimental Study
JMIR Aging 2025;8:e67889
URL: https://aging.jmir.org/2025/1/e67889
doi:10.2196/67889

Implementation of New Technologies in an Aged Care Social Day Program: Mixed Methods Evaluation

Dannielle Post, PhD; Kathleen Whitson, BPsychSc (Hons); Gaynor Parfitt, Prof, PhD

Alliance for Research in Exercise, Nutrition, and Activity, Allied Health and Human Performance, University of South Australia, City East Campus, Frome Road, GPO Box 2471, Adelaide, Australia

Corresponding Author:
Gaynor Parfitt, Prof, PhD
Alliance for Research in Exercise, Nutrition, and Activity, Allied Health and Human Performance, University of South Australia, City East Campus, Frome Road, GPO Box 2471, Adelaide, Australia

Abstract

Background: Australia's aging population is looking to age in place, accessing care alternatives external to the traditional model of residential aged care facilities. This evaluation is situated in a Social Day Program, delivered by an aged care organization. It is designed to cater for people living with dementia, located in an environment equipped with new technologies including age-specific interactive computer gaming, social robots, sensory stimulation, and virtual reality. The technologies are designed to support older adults, enabling them to stay connected and maintain physical and cognitive functioning, independence, and quality of life.

Objective: This project aimed to undertake a multifaceted evaluation of the implementation of the new technologies, including an exploration of the barriers and enablers to uptake. The key issue is how to enhance the potential for optimizing the use of these technologies in the Social Day Program environment, to help inform decision-making regarding the implementation of these technologies at the organization's other sites, and future investment in such technologies by aged care organizations generally.

Methods: Observation of technology use within the organization was conducted over a 16-week period. Surveys and semistructured interviews were used to collect information from staff related to their experiences with the technology. Thematic analysis was used to analyze the interviews. Data were triangulated across the sample.

Results: Forty-eight observation periods were completed, totaling 126.5 observation hours. Technology use by clients was observed on 24 occasions, for 22 (17.4% of the observation time) hours. Nineteen staff completed surveys. Nearly three-quarters (n=14) of the staff perceived there to be barriers to the clients' use of technology, and 18 (95%) staff reported that they assisted clients to use the technology. Ten (53%) staff reported receiving training to use the technology and feeling confident in their knowledge of the technology to assist clients in using it. Twelve staff members participated in an interview. Key themes identified from the interview data were: technology has potential but is not for everyone, incorporating the subtheme technology as a placation tool, staff knowledge and confidence, and technology functionality and support.

Conclusions: This evaluation identified that technology was not being used for the purposes of enrichment or experience enhancement, nor extensively. Multiple barriers to the implementation and sustained use of the technology items were identified. Recommendations to improve implementation and promote sustained use of technology, based on the findings of this evaluation and evidence from the literature, may apply to other organizations seeking to implement these technologies in similar programs.

(JMIR Aging 2025;8:e60297) doi:10.2196/60297

KEYWORDS

aged care; older adults; interactive robots; social engagement; evaluation; geriatric; robot; day program; perception

Introduction

Background

With a shift to consumer-directed care, and Australia's aging population looking to age in place and access a range of supported care alternatives outside the traditional model of residential aged care facilities, there are a lot of new offerings for consumers. Often funded through the Australian government–supported Aged Care Community Packages, these offerings can include in-home care and home support services or colocated aging-specific services in centralized hubs. Services include Dietetics, Allied Health Group Programs, Exercise Physiology, Gym and Physical Activity, Massage, Music Therapy, Occupational Therapy, Physiotherapy, Podiatry, Speech Pathology, Gerontology, and Social Day Programs.

ECH Inc (hereinafter referred to as the organization) is a South Australian retirement living and in-home care services provider established in 1964 and based in the state of South Australia. The "profit for purpose" organization's mission is "helping

people to live confidently and independently, and to get more out of life" [1]. This project was situated in one of the organization's Social Day Programs, which occur at 4 centers across metropolitan South Australia. The Social Day Program involved in this evaluation is situated within a proof-of-concept Care Hotel that was purpose-designed to provide care givers respite options with 8 short-stay suites for people living with dementia or who may be recovering from surgery. These Social Day Programs, provided on weekdays between 9 AM and 3 PM, are designed to cater for community-dwelling people living with dementia or memory loss and are located in environments where the clients have space to engage socially with others or have supported, quiet alone time. Clients can be dropped off at and picked up from the Social Day Program by family members or transported to and from the Social Day Program by a bus provided by the organization. Program staff are trained to ensure that each client's interests and abilities are met, with the provision of activities that are enjoyable and meaningful. Participation in the Social Day Program is also available to Care Hotel guests, as is access to the space outside the Social Day Program hours (evenings and weekends). This space is equipped with new technologies that are designed to support older adults as they age, enabling them to stay connected and maintain physical and cognitive functioning, independence, and quality of life [2-5]. These technologies are varied and includes an age-specific interactive computer gaming console, "Obie" that projects interactive games onto a surface (table or floor) to encourage active play; a relaxation chair in a sensory stimulation room; virtual reality (VR) headsets (the Odysee system, with virtual experiences including travel, museums, and hot air ballooning); white noise bubble tubes (tubes with color changing lights, filled with bubbles, that are intended to provide visual distraction and calm); 2 robotic dolls; and a robotic sensory cat.

Systematic reviews support the effectiveness of these technologies, with interactive computer gaming improving physical and psychological functioning [6]; interactions with social robots positively affecting agitation, anxiety and loneliness, medication consumption, and quality of life [5]; and VR technology improving physical health outcomes [7] and cognition, memory, and depression [8] for older adults. The implementation of such technologies in the aged care environment is not without its challenges. Other studies have identified organizational factors, such as funding and staff engagement or knowledge, client-specific issues such as frailty, dementia, or limited prior exposure to technology, inappropriate or unsuitable technologies, or resistance from family members [9]. Similar barriers were identified in a systematic review of VR technology, with the addition of the potential for cybersickness during use also perceived to be a barrier [10]. Research focused on social and assistive robots has identified concerns around the potential for infantilization or loss of dignity for older adults who use them [11]. The items chosen by the organization were selected due to the evidence supporting their effectiveness and with the intention of supporting engagement and enrichment for their clients. Providing a variety of technologies, rather than a single type of technology, was intended to broaden the appeal of technology to clients.

This Study

The organization requested an extensive review of the use of these technologies in this unique setting. This research aimed to undertake a multifaceted evaluation of the implementation of technologies at the Social Day Program, including an exploration of the barriers and enablers to the use of technology within the Social Day Program. The key issue, and the focus of this paper, is identifying how to enhance the potential for optimizing the use of these technologies at the organization, making recommendations to help inform decision-making regarding the implementation of these technologies at other sites, and future investment in new technologies for the organization's clients.

The evaluation was intended to answer the following research questions:

1. What is the utility of these technologies?
2. What are the barriers and enablers to the uptake of these technologies?
3. What is the impact (physical, psychosocial, and quality of life) of technology use experienced by the clients, as perceived by the staff?

Methods

Study Design

Observation of technology use within the Social Day Program was conducted over 16 weeks by a research assistant (RA), with periodic support from undergraduate physiotherapy students completing a health promotion placement with the evaluation team. Surveys and semistructured interviews were used to collect information from staff related to their experiences with the technology. Data were triangulated across the sample, with observation, survey, and interview data integrated and compared to establish alignment and contrast among the findings of each method.

Participants

Participants were staff working in the organization's Care Hotel or Social Day Program. Staff provided written informed consent to participate in the evaluation. For the observation component, participants included clients of the Social Day Program and clients of the Care Hotel who were accessing the Social Day Care space.

Observation of Technology Use

Observation of the use of technology by staff, Social Day Program clients, and Care Hotel guests using the Social Day Program space was undertaken between July and November 2022. Observation periods of 4 hours each, either morning and overlapping early afternoon or late morning until the end of the day, occurred during the Social Day Program sessions across the week, but also after hours and on weekends to ensure capture of use of technology outside of typical program activity periods. An observation checklist was used to capture data related to technology-user type (ie, client, family member or carer, staff, or other), technology item, duration of use, concurrent use of an item by multiple participants, support or assistance provided, emotional impacts, communication, physical impacts, and other

notable information. This checklist was developed by the evaluation project team in consultation with the Stakeholder Advisory Group and piloted by the RA and placement student observers, with checklist items modified as required throughout piloting.

Surveys

Consenting staff completed an anonymous electronic survey (Qualtrics). Survey items were initially drafted by the evaluation team and refined using a co-design approach in collaboration with the Stakeholder Advisory Group. The Stakeholder Advisory Group was comprised of members of the organization and evaluation project team and a representative from each of the 3 stakeholder-participant groups (clients, family members or carers, and staff of the Social Day Program) and was formed to ensure consumer awareness and the suitability and feasibility of evaluation methods. The staff survey was intended to identify the perceived implications for staff involved in implementing the technology and supporting clients to use the technology.

Survey questions collected information about participant demographics, staff assistance with technology items, length of time staff assisted clients with the technology, and most- and least-assisted technology items. Perception of impact were collected across 6 domains relative to each of the items of technology: social engagement, cognitive awareness, communication, mood state, activities of daily living, and physical mobility. Responses were on a 4-point Likert scale, with responses being: don‘t know, no improvement, some improvement, and a lot of improvement. Additionally, there were questions relating to whether staff perceived that clients disliked any technology items; barriers to clients using the technology; staff training; staff confidence when supporting clients; and barriers to staff supporting clients with the technology items.

Interviews

Semistructured, once-off, one-on-one interviews were conducted by 1 of 2 University of South Australia RAs, with consenting staff. Interview questions were initially drafted by members of the evaluation project team and refined and co-designed in collaboration with the Stakeholder Advisory Group (Multimedia Appendix 1). Interviews were conducted either in person at the Social Day Program or over the phone. All interviews were recorded and then transcribed verbatim.

Data Analysis

Observation and survey data are reported as counts or counts and percentage responses as relevant. Content analysis of open-ended survey questions was intended; however, an insufficient number of responses was provided, and a descriptive approach to reporting the responses was instead used. Reflexive Thematic Analysis [12,13] was used to analyze interview data, which were coded independently by a member of the evaluation team. Semantic codes and candidate themes were reviewed and discussed among the team across multiple iterations before the final themes were decided. We applied a constructionist epistemology, with an experiential lens on the data, with the intention of highlighting the experiences and perspectives of staff related to the use of technology in the Social Day Program environment. While our approach was predominantly inductive, in that we sought to represent the meaning of the information as it was communicated by the staff, elements of a deductive approach were applied to ensure that the established themes aligned with the research questions, for example, what the staff perceived to be barriers to and enablers of technology use in this setting.

Ethical Considerations

This evaluation research was approved by University of South Australia’s Human Research Ethics Committee (204457). A waiver of consent was approved for the purpose of observing technology use in the Social Day Program. This was on the proviso that before commencing each observation period, staff and the evaluation project team announced to clients during Circle Time—an activity during which information is relayed to clients and their family members—that an observation period was about to begin. This enabled clients to leave the room if they did not wish to participate in this part of the evaluation project. Participants remained anonymous, in that information that might identify them was not recorded per their use of technology, and they were not asked to modify their behavior for the evaluation. Participants provided written informed consent for use of their survey data. Written informed consent was also provided by interview participants. Each interview participant received an honorarium (AUS $30 gift card; US $19.43) to acknowledge their contribution to the project. Data are anonymized and no information is provided that would identify participants.

Results

Participants

Thirty-one staff were approached to participate in the evaluation project; 19 based in the Care Hotel, and 12 staff based in the Social Day Program. Nineteen (61.3%) staff completed surveys (9 Care Hotel and 10 Social Day Program). Ten (32.3%) of the 31 staff approached (3 Care Hotel and 7 Social Day Program) provided written informed consent to participate in the interview; all 10 interviews were completed.

Observation of Technology Use

Forty-eight observation periods were completed, totaling 126.5 observation hours. Technology use by clients was observed on 24 occasions, for 22 (17.4% of the observation time) hours. The relaxation chair and the robotic cat were the most frequently used technology items. The relaxation chair was used 7 times for a total of 9 hours and 36 minutes, and the robotic cat was used 7 times for a total of 3 hours and 14 minutes. Conversely, the Obie interactive games table and VR headsets were the least frequently used technology items (once for 64 min and once for 39 min, respectively). When technology was used, clients were mostly observed socially engaging and communicating with staff and other clients, with some clients also observed speaking to the robotic cat and dolls.

Surveys

Of the 19 staff members who consented and responded to the staff survey, 10 (53%) worked in the Social Day Program and

9 (47%) were Care Hotel staff. Ninety-five percent (n=18) of staff reported that they helped or supported clients in using the technology items. Social Day Program staff assisted clients more than Care Hotel staff across all technology items except for the relaxation chair. Furthermore, all 9 Social Day Program staff indicated they assisted clients with the Obie interactive games table and robotic cat. Seventeen of the nineteen (89%) staff members nominated the VR headsets as the technology item they engaged with the least. Staff reported that on average, clients engaged with the technology items from 10-minutes to more than 60-minutes. This finding is reflected in the observation data, with the average time clients spent using technology items ranging from 32-minutes to 96-minutes.

Improvements in social engagement, cognitive awareness, communication, mood state, activities of daily living, and physical mobility were seen to some extent across each of the technologies, although not for all items (Table 1).

Table . Perceived improvements associated with technology use.

Technology item and domain		Don't know, n (%)	No improvement, n (%)	Some or a lot of improvement, n (%)
Obie ICG[a] table				
	Social engagement	1 (5)	3 (16)	15 (79)
	Cognitive awareness	3 (16)	4 (21)	12 (64)
	Communication	1 (5)	2 (11)	16 (84)
	Mood state	1 (5)	0 (0)	18 (95)
	Activities of daily living	2 (11)	9 (47)	8 (42)
	Physical mobility	0 (0)	8 (42)	11 (58)
VR[b] headsets				
	Social engagement	10 (53)	4 (21)	5 (26)
	Cognitive awareness	12 (63)	5 (26)	2 (11)
	Communication	12 (63)	4 (21)	3 (16)
	Mood state	10 (53)	4 (21)	5 (26)
	Activities of daily living	11 (58)	7 (37)	1 (5)
	Physical mobility	12 (63)	6 (32)	1 (5)
Robotic cats				
	Social engagement	2 (11)	1 (5)	16 (84)
	Cognitive awareness	2 (11)	5 (26)	12 (63)
	Communication	2 (11)	2 (11)	15 (79)
	Mood state	2 (11)	0 (0)	17 (90)
	Activities of daily living	3 (16)	11 (58)	5 (26)
	Physical mobility	3 (16)	8 (42)	8 (42)
Robotic dolls				
	Social engagement	1 (5)	1 (5)	17 (90)
	Cognitive awareness	1 (5)	4 (21)	14 (74)
	Communication	3 (16)	1 (5)	15 (79)
	Mood state	2 (11)	0 (0)	17 (89)
	Activities of daily living	5 (26)	8 (42)	6 (32)
	Physical mobility	4 (21)	6 (32)	9 (47)
Relaxation chair				
	Social engagement	2 (11)	7 (37)	10 (52)
	Cognitive awareness	4 (21)	5 (26)	10 (53)
	Communication	3 (16)	5 (26)	11 (58)
	Mood state	1 (5)	0 (0)	18 (95)
	Activities of daily living	3 (16)	6 (32)	10 (53)
	Physical mobility	2 (11)	6 (32)	11 (58)

[a]ICG: interactive computer gaming.

[b]VR: virtual reality.

Nearly three-quarters (n=14) of the responding staff perceived there to be barriers to the clients' use of technology, with equal numbers responding "yes" when compared by staff type. Ten (53%) of the responding staff had received training. Ten (53%) staff indicated they felt confident when assisting clients to use the technology; however, ten (53%) staff also felt that barriers existed that inhibited them from assisting the clients to use the technology (5 each from the Social Day Program and Care Hotel).

Key concepts raised by staff in their responses to the open-ended questions about their perceptions of technology use within the Social Day Program related to why clients may dislike technology, the barriers to technology use by clients, factors related to staff confidence in supporting clients to use technology, and the barriers to staff supporting clients to use technology.

Staff perceived that clients may dislike the technology because they have health or cognition issues that limit their ability to engage with technology; due to negative perceptions or stigmatization by other clients of people who use the robotic doll or cat; that clients have no interest in technology or use of technology; and that clients did not like particular items, finding them to be noisy or creepy (ie, the robotic doll). These aspects were also considered to be barriers to clients' use of technology, with the addition of staff knowledge and confidence in the use of technology, and the time impact on staff to assist clients in using the technology.

Interviews

Interviews were conducted with 12 staff members to assess the use of technology within the Social Day Program, and the enablers of and barriers to technology use. Interviews ranged between 18 and 44 minutes. The three themes developed during the analysis of the interview data: (1) technology has potential but is not for everyone, with the subtheme technology as a placation tool; (2) staff knowledge and confidence in using technology; and (3) technology functionality and support, are based on our interpretation of how staff experienced technology use in the Social Day Program and the factors that are likely to be enablers or barriers that influence the adoption of technology and sustainability of technology use in this environment. These themes address the research questions that form the basis for the evaluation. Quotes from participants are used as illustrative examples of the themes established.

Technology Has Potential but it Is not for Everyone

This theme reflects the value and appropriateness that the staff perceive in technology use for their clients, assuming who might be likely to benefit from technology use and, conversely, for whom technology is not for. This theme incorporates the safety and personal considerations that were voiced as being needed when using technology. There were conflicting perspectives related to the use of technology within the Social Day Program. The potential benefits of the technology were raised often, including increased social engagement through the use of interactive games or being able to placate agitated clients with the robotic cat or time in the relaxation chair (a subtheme). While there was some concern about the cognitive benefits of the technologies, that the clients had to sit for prolonged periods to use some of them and therefore were being sedentary just as they would be at home (staff 14), and that some items had the potential to cause "sensory overload" (staff 2), others thought that the technology promoted social interaction, could "enrich the time clients spent at the Social Day Program" (staff 24), and could benefit future clients, in particular those who were tech-savvy. For one staff member, their initial perception that technology would "not be a great idea" (staff 1) changed positively when they saw the impact of technology use for some clients, for example, the connections made with the robotic cat. Similar observations of such connections are reflected by the following comment from a staff member.

> *I wasn't sure about the babies and the cat. I thought maybe that was a bit babyish...It's funny to watch how they [clients] perceive that and how quick they are to defend the cat, whereas sometimes you wouldn't even get a couple of words out of them, but they've got to make sure the cat is okay.* [Staff 7]

Generational factors, in that the client demographic needed assistance as they were not tech-savvy, or not familiar with, or not interested in technology, situated some clients in the "technology is not for everyone" group, in the eyes of the staff. These factors were often raised as a possible barrier to greater use of the technology in the Social Day Program. Future generations were viewed to be more likely to be receptive to technology, as evidenced by the following comments:

> *I think it's a double-edged sword and it has the potential to be brilliant, however unfortunately with that particular generation, because they were not particularly technologically minded or advanced, they are somewhat overwhelmed by it perhaps.* [Staff 6]

> *Look, we do have to help them most times just because they are technologies and it's not their era. Today's kids will be great in the future for these things, but these people were not brought up with these technologies.* [Staff 4]

While there was acknowledgment of the potential benefits of technology use, it was clear that staff were wary of the possible downfalls associated with it, demonstrating concern around physical and psychological safety for clients with frailty or cognition issues, or for whom memories of past traumas were triggered by using technology. Examples of this included clients who had worked in child abuse situations and had experienced miscarriages or the death of infants being exposed to the robotic dolls.

> *Some clients are really not suitable for the dolls and some aren't suitable for the animals. It does depend, because they've got either histories of – we've got some clients that have worked in child abuse situations. They're not good with the dolls, because it really upsets them. Knowing that background as well...we do the assessments and we ask the questions. But it's good knowing those things.* [Staff 4]

The juxtaposition between the potential benefit of technology use and its appropriateness was evident in the commentary around the use of the VR equipment. There was a common thread that the VR equipment could be the most beneficial for clients of all of the technologies introduced to the Social Day Program, allowing clients to reminisce about places they may have previously lived in or visited and activities previously performed, such as flying planes. For others, the VR was a source of concern; for example, clients with frailty issues may be scared to use it due to a fear of falling, and those with cognitive issues may be confused by what they are seeing.

> *I think that the VR would definitely be useful for some clients who may enjoy it, a lot of more cognitive ones I think, because people who have low cognition may get very frightened of some of the things they see.* [Staff 1]

Ensuring the safety of clients while they were using technology by having measures in place to assess physical capacity to use the VR, for example, or having background information that enabled identification of potential trauma triggers to avoid clients being exposed to them, was raised, with some staff indicating processes were in place within the organization to accommodate this.

> *We've got a safety test before we just chuck anybody on it [VR], because some people do have the same thing as me with the nausea and dizziness and stuff. There is a possibility of falling down.* [Staff 4]

Combined, not being tech-savvy, being averse to technology, or perceived to be at risk of physical or psychological injury due to technology use cemented the concept that technology is not for everyone involved in the Social Day Program.

Technology as a Placation Tool

Incorporated within this theme is the subtheme of technology as a placation tool. We initially wrestled with whether the use of technology in this way was situated as a potential benefit of technology use or a standalone theme; however, throughout the data, the utility of technology as a method of placation for clients, and therefore, it being perceived as a potential benefit, was evident. It seemed that while this was not an intended application of technology within the Social Day Program, placating agitated clients with items such as the robotic cat or time in the relaxation chair was rationalized by staff to lead to enriched experiences for those clients, as well as other clients in the program, whose experience was no longer being impinged upon by an agitated client. As a by-product, the benefit was extended to family members of the agitated clients who did not have to pick clients up from the Social Day Program earlier than intended, as described in the following comment:

> *It makes them extend their stay here in the day program. Rather than having to be picked up at lunchtime because their day is done and they've had enough stimulation for the day, they can be in there [the sensory room] for half an hour to an hour, come back out and enjoy the rest of the day 'til 3:30 [PM].* [Staff 4]

Knowledge and Confidence in Using Technology

This theme encapsulates concepts related to staff knowledge of the capabilities of the various pieces of technology and how to use them, as well as their confidence in assisting clients to use the technology, as a key driver in the use of technology in the Social Day Program. Further, these aspects are likely to influence whether technology continues to be used in the program. It was apparent across the data that training was important to staff; however, most reported that they had received little to no training in the use and maintenance of the technology. Further, there were perceptions that some of the technology required little training, such as the robotic doll and cat, whereas other types of technology, such as the VR or Obie, required more training of staff. This may have contributed to lower usage of these technologies as evidenced through observation, survey, and interview data, as may have confidence in using the technology. Perceptions of the relationship between training and confidence to use technology are contrasted. In some cases, staff reported that more training in the use of the available technology would increase their confidence to use the technology, for example:

> *I would like more training and capacity. I can do basic things, but I'm not confident using things...Yeah, I think I don't feel really confident with it. I'd like to be. I'd like to learn to do it.* [Staff 9]

Others felt confident to use the technology, irrespective of the level of training they had received. It was suggested that increasing confidence in the use of technology may increase the enthusiasm staff have for technology use, and subsequently, increase the regularity of technology use in the programs at the organization and assist in better tailoring technology use to individual clients.

Technology Functionality and Support

This theme reflects the importance of having available technology that is functional, as well as accessible support for staff in cases when the technology was not working. This would seem obvious in any environment where technology such as that implemented was being used; however, it was frequently raised by interviewees that technical and maintenance issues with the technology were common. These are clear barriers to the use of technology in the Social Day Program and are likely to influence the willingness of staff to use the technology or assist clients to use it. An example of this related to the VR technology, which staff reported had rarely been in working condition since its introduction. While some staff reported helping other staff to use the technology, obtaining higher-level support to repair the equipment had proved problematic, as illustrated by this comment from a staff member.

> *[If] I had a problem with something over the weekend let's say, I would have absolutely no idea who to call to get that support and I don't even think that the support necessarily is available. Therefore, from that perspective those clients would simply miss out until someone was able to be contacted to fix it. It would be a too bad, so sad kind of concept, which is not cool really.* [Staff 6]

Discussion

Principal Findings

This project aimed to undertake a multifaceted evaluation of the implementation of technologies at a Social Day Program to help inform decision-making regarding their implementation at the organization's other sites. The evaluation included (1) observation of use, (2) surveys of staff members' perceptions of enablers and barriers to use of the technologies, and (3) interviews with staff to further explore their experiences in use of technology and enablers and barriers to implementation and use of the technologies. The observation, staff survey, and

interviews provide a rich dataset to enable the evaluation. Observation over an extensive period identified limited use of technology, further verified by the survey data, with potential reasons for this, including barriers to and enablers of technology use identified in survey data and thematic analysis of interview data.

When technology was used it was predominantly used during the Social Day Program and by Social Day Program clients rather than after hours by the Care Hotel guests. One reason for this may be that guests in the Care Hotel might be recuperating from medical procedures, rather than solely attending for respite purposes. As such, these guests may not be able to engage, nor be interested in engaging, in technology use at this time. Alternatively, the technology located in the guests' Care Hotel rooms, such as an interactive television and tablet, which were not the focus of this evaluation, may have negated the need to access the technology located in the Social Day Program. For example, guests could relax and listen to music in their hotel rooms, and therefore would not need to access the sensory room and relaxation chair.

While items such as the robotic cat have been shown to be popular, a large portion of the use appears to be as a means of placating clients who were agitated or were disturbing other clients. This is not to say that the potential benefits of technology use at the organization were not recognized by some staff; however, encouragement of use may not be aligned with the original intention behind the implementation of the technology. Staff highlighted occasions when the use of the robotic cat by clients enhanced interaction between those clients, other clients, and staff. While mixed findings have been reported concerning the use of robotics in other aged care environments, some studies have shown that robotic animals such as cats and dogs can be conversation starters and contribute to enhanced engagement for older adults and people living with dementia [14,15].

In other cases, staff spoke of the opportunity that the sensory room provided for clients in being able to support an extended period of attendance at the Social Day Program, and in their eyes, held this as a potential benefit of technology use. Clients who may normally need to be picked up early, as they found the activities of the Social Day Program tiring or overwhelming, could instead remove themselves from the activities and take some time-out to relax in the sensory room. Being able to do this meant that these clients could spend longer days at the Social Day Program, rather than having to go home early. This has flow on benefits for their family carers, who, as a result, have a longer period on that day for respite. Research supports a reduction in care-related stressors for the carer during respite, and that the longer the respite duration, the greater the benefit to the carer [16]. Each of these factors can be highlighted to staff as potential benefits as a means of supporting implementation and greater use of the technology.

Harnessing the staff members' own perceptions of the emotional, social, cognitive, and physical benefits of technology use that were identified through this evaluation may be one approach to facilitating successful implementation. Observational, survey, and interview data indicated that technologies were perceived to positively impact the emotional state of clients. This was not only the calming effect reported and that is acknowledged in the literature [17], but also included observed happiness and enjoyment, as well as the social engagement that occurred among groups of Social Day Program clients as they used the Obie and robotic cats, for example.

Safety—both physical and psychological—was a prominent consideration for staff. Caution was raised about the risk of falls with some of the technology, in particular VR. The use of VR has been evaluated in other aged care environments and has been shown to have both positive and negative impacts. While it has the potential to engage and enrich the experience of the organization's clients by providing them access to activities, experiences, and environments that may no longer be physically accessible to them, VR may not be acceptable to, nor appropriate for, all people living with dementia [18]. Raised in the survey and interview data, these safety concerns extended to the possibility of psychological trauma for clients exposed to technology such as the robotic cats, categorizing them as part of the "technology is not for everyone" group. Concerns were raised about the stigma associated with people who use technologies such as the robotic cat or doll, because they are "not real." "Infantilization" of older adults living with dementia and their loss of dignity through the use of robotic toys, which they perceive to be real, has been explored in other research [17], and in some cases, it is the family members of the older adult holding this perception [19]. Strategies suggested to minimize this and maintain dignity when items such as robotic toys are used to engage people include creating an obvious environment of "play" [20].

Barriers to, and Enablers of, Technology Uptake and Use

As evidenced through this evaluation, several factors act as barriers to and enablers of technology use within the Social Day Program. The key factors, reflected by the developed themes, relate to staff training in, and knowledge of, technology operation and features; technology functionality and the availability of real-time technology support for staff; and perceptions of a generational influence on technology use. All of the barriers—and subsequently the enablers—to technology use identified in this evaluation were common to another recent Australian evaluation of technology use in aged care [9], suggesting they are not specific to this particular organization, but more reflective of the aged care environment generally.

Staff awareness and understanding of the underlying purpose of technology use for this population, beyond it being used as a behavior management tool, may also act as an enabler of greater use of technology within the organization. This includes opportunities for technology to promote "meaningful engagement" and social interaction [19], particularly, the technology should facilitate social interaction among those present and not be a replacement for it [9]. However, technologies that require staff to invest more of their already limited time to support clients in using them may also be less likely to be encouraged within the Social Day Program.

"Visibility" of the technology was considered by some staff to be an enabler of technology use; however, it was not raised sufficiently to warrant inclusion as a theme. Staff suggested

that technology items that are visible and easily accessible by staff and clients are more likely to be used regularly. Further, staff members who regularly use the technology and encourage clients to do so have the potential to encourage other staff members to do the same. As such, increased visibility of the technology has the potential to encourage more use within the organization.

Other potential barriers to the use of technology relate to perceptions that the current generation of clients at the Social Day Program are not "tech-savvy," do not like technology, or just have no interest in using it. Interviews identified that the clients were perceived not to have used much technology across their lifetime and that technology is likely to be more acceptable to future generations of clients who have grown up with and are familiar with technology. Staff reported observing positive impact related to social engagement, communication, and mood state, suggesting that some clients did enjoy using the technology. Enjoyment and acceptability of technology by older adults would seem to be obvious in whether technology is used and have been shown to be perceived by older adults as relevant in the successful implementation of technology to support ageing in place [20].

Sustainability

Targeting the enablers of technology use identified in this evaluation will contribute to ongoing technology use in the organization's Social Day Program. Taking a person-centered approach to technology use, by modifying organizational approaches to understanding the interests of clients, the physical and cognitive capabilities of clients, knowledge of potential triggers for negative or trauma responses associated with technology use; as well as technology that is maintained in good working order will also be a driver of sustainability.

Limitations

A limitation of this evaluation was having little data from clients and their family members, and therefore not being able to make their voices prominent in this evaluation. To an extent, this is reflective of other evaluations that have been undertaken in the aged care environment. In some cases, family members may not be involved in the lives of their guest or client family member. In others, the only respite family members may have is when the guest or client is at the Social Day Program. This may see family members not wanting to use that time for activities such as this evaluation.

In cases where guests or clients are living with cognitive decline or dementia, it was difficult to engage them in the interview process. However, it is important to ensure that people living with cognitive decline or dementia are included in projects about them, in particular, in the consultation and co-design process [21], which we were able to achieve through having clients and family members in the stakeholder group. While there were several instances in which staff provided negative feedback related to the technology not being operational, with limited support provided, it is also possible that other staff were reluctant to provide negative perspectives, which may introduce some bias to the findings.

Recommendations

While the use of technology may benefit this population, the potential benefits have not been fully realized at the organization due to several implementation barriers: there is a lack of staff knowledge, confidence in use, and training relevant to the operation of the different technologies, the technology has not always been operational and ongoing, accessible support is not available. There is a perception that technology has benefits, although the use of technology may not always be as intended by the organization, however, this is contrasted by the expressed view that technology is not for everyone and there are physical and psychological risks involved in use of technology in the Social Day Program population, many of whom live with cognitive decline. By addressing the barriers identified in this evaluation and strengthening the enablers per the following recommendations, the potential for clients to benefit across the domains could be enhanced.

Based on the results from the evaluation and considering the evidence from the literature, the following recommendations may enhance the implementation of technology in programs similar to the one described here:

- Increase staff knowledge of technology and its potential benefits for older adults living with cognitive decline, through the provision of more structured and regular training, a summary information sheet of project findings, quick refresher sessions, instruction sheets, and time to "play" and practice with the technology. This needs to include not only how to operate the technology and trouble shoot common problems or issues, but also provide an understanding of all the features and options (to enable tailoring to client needs or capabilities) and how to use the technology to best benefit the clients.
- Develop an information technology maintenance and support process that is resourced to ensure the equipment is functional at all times and enables staff to access real-time information technology support to address operational or technical issues. This should also include the development of comprehensive but user-friendly instructions to assist with understanding the features of the technology, operating the technology, and troubleshooting.
- Ensure that appropriate measures are in place to establish client physical and psychological safety, so that they are not exposed to technology that may be inappropriate for their circumstances.
- Increase the visibility and accessibility of the technology, so that it is at the forefront of staff's, clients' and their family members' or carers' thoughts as an activity. This may be through pictural posters placed around the room as reminders that the technology is there, with simple step by step instructions placed near technology or by leaving the technology in easy-to-access locations.
- Consider the placement of technology within the space.

Conclusions

This evaluation of technology use in a Social Day Program has identified that technology is not being used extensively, nor is it being used for the purposes of enrichment or experience enhancement. Multiple barriers to implementation and sustained

use of the technology items have been identified, spanning perceptions of clients' preference or low ability, to a lack of staff training and knowledge to support adequate use of the technology. Recommendations to improve implementation and promote sustained use of technology, based on the findings of this evaluation and evidence from the literature, may apply to other organizations seeking to implement these technologies in similar programs.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guiding questions: family member or carer.
[DOCX File, 22 KB - aging_v8i1e60297_app1.docx]

References

1. About us. ECH. 2024. URL: https://www.ech.asn.au/about-us [accessed 2025-05-02]
2. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020 Aug 11;3(2):e17286. [doi: 10.2196/17286] [Medline: 32780020]
3. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: a systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
4. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults' social isolation. Comput Human Behav 2016 Oct;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
5. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
6. Bleakley CM, Charles D, Porter-Armstrong A, McNeill MDJ, McDonough SM, McCormack B. Gaming for health: a systematic review of the physical and cognitive effects of interactive computer games in older adults. J Appl Gerontol 2015 Apr;34(3):NP166-NP189. [doi: 10.1177/0733464812470747] [Medline: 24652863]
7. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
8. Yen HY, Chiu HL. Virtual reality exergames for improving older adults' cognition and depression: a systematic review and meta-analysis of randomized control trials. J Am Med Dir Assoc 2021 May;22(5):995-1002. [doi: 10.1016/j.jamda.2021.03.009] [Medline: 33812843]
9. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162. [doi: 10.2196/31162] [Medline: 34975014]
10. Kouijzer M, Kip H, Bouman YHA, Kelders SM. Implementation of virtual reality in healthcare: a scoping review on the implementation process of virtual reality in various healthcare settings. Implement Sci Commun 2023 Jun 16;4(1):67. [doi: 10.1186/s43058-023-00442-2] [Medline: 37328858]
11. Rigaud AS, Dacunha S, Harzo C, et al. Implementation of socially assistive robots in geriatric care institutions: healthcare professionals' perspectives and identification of facilitating factors and barriers. J Rehabil Assist Technol Eng 2024;11:20556683241284765. [doi: 10.1177/20556683241284765] [Medline: 39324049]
12. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
13. Braun V, Clarke V. Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Couns Psychother Res 2021;21(1):37-47. [doi: 10.1002/capr.12360]
14. Campbell A. Dementia care: could animal robots benefit residents? Nurs Resid Care 2011 Dec;13(12):602-604. [doi: 10.12968/nrec.2011.13.12.602]
15. Naganuma M, Ohkubo E, Kato N. Use of robotic pets in providing stimulation for nursing home residents with dementia. Stud Health Technol Inform 2015;217(1009-12):1009-1012. [doi: 10.3233/978-1-61499-566-1-1009] [Medline: 26294602]
16. Zarit SH, Kim K, Femia EE, Almeida DM, Savla J, Molenaar PCM. Effects of adult day care on daily stress of caregivers: a within-person approach. J Gerontol B Psychol Sci Soc Sci 2011 Sep;66(5):538-546. [doi: 10.1093/geronb/gbr030] [Medline: 21642593]
17. Koh WQ, Ang FXH, Casey D. Impacts of low-cost robotic pets for older adults and people with dementia: scoping review. JMIR Rehabil Assist Technol 2021 Feb 12;8(1):e25340. [doi: 10.2196/25340] [Medline: 33497349]
18. Baker S, Waycott J, Robertson E, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020 May;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]

19. Neal I, du Toit SHJ, Lovarini M. The use of technology to promote meaningful engagement for adults with dementia in residential aged care: a scoping review. Int Psychogeriatr 2020 Aug;32(8):913-935. [doi: 10.1017/S1041610219001388] [Medline: 31547900]
20. Peek STM, Wouters EJM, Luijkx KG, Vrijhoef HJM. What it takes to successfully implement technology for aging in place: focus groups with stakeholders. J Med Internet Res 2016 May 3;18(5):e98. [doi: 10.2196/jmir.5253] [Medline: 27143097]
21. Goeman DP, Corlis M, Swaffer K, et al. Partnering with people with dementia and their care partners, aged care service experts, policymakers and academics: a co-design process. Australas J Ageing 2019 Sep;38 Suppl 2(S2):53-58. [doi: 10.1111/ajag.12635] [Medline: 31496067]

Abbreviations

RA: research assistant
VR: virtual reality

Edited by R Yang; submitted 07.05.24; peer-reviewed by A Haynes, J Waycott, R Kelly; revised version received 25.03.25; accepted 29.03.25; published 12.05.25.

Please cite as:
Post D, Whitson K, Parfitt G
Implementation of New Technologies in an Aged Care Social Day Program: Mixed Methods Evaluation
JMIR Aging 2025;8:e60297
URL: https://aging.jmir.org/2025/1/e60297
doi:10.2196/60297

Unveiling the Frailty Spatial Patterns Among Chilean Older Persons by Exploring Sociodemographic and Urbanistic Influences Based on Geographic Information Systems: Cross-Sectional Study

Yony Ormazábal[1*], MSc; Diego Arauna[2*], PhD; Juan Carlos Cantillana[3], MSci; Iván Palomo[2], PhD; Eduardo Fuentes[2], PhD; Carlos Mena[1], PhD

[1]Longevity Center VITALIS, Faculty of Economics and Business, University of Talca, Universidad de Talca, Avenida Lircay S/N, Talca, Chile

[2]Thrombosis Research Center, Department of Clinical Biochemistry and Immunohematology, Faculty of Health Sciences, Interuniversity Center of Healthy Aging (CIES), Longevity Center VITALIS, University of Talca, Talca, Chile

[3]Facultad de Administración y Economía, Universidad Tecnológica Metropolitana, Santiago, Chile

[*]these authors contributed equally

Corresponding Author:
Carlos Mena, PhD
Longevity Center VITALIS, Faculty of Economics and Business, University of Talca, Universidad de Talca, Avenida Lircay S/N, Talca, Chile

Abstract

Background: Frailty syndrome increases the vulnerability of older adults. The growing proportion of older adults highlights the need to better understand the factors contributing to the prevalence of frailty. Current evidence suggests that geomatic tools integrating geolocation can provide valuable information for implementing preventive measures by enhancing the urban physical environment.

Objective: The aim of this study was to analyze the relationship between various elements of the urban physical environment and the level of frailty syndrome in older Chilean people.

Methods: A cohort of 251 adults aged 65 years or older from Talca City, Chile, underwent comprehensive medical assessments and were geographically mapped within a Geographic Information Systems database. Frailty was determined using the Fried frailty criteria. The spatial analysis of the frailty was conducted in conjunction with layers depicting urban physical facilities within the city, including vegetables and fruit shops, senior centers or communities, pharmacies, emergency health centers, main squares and parks, family or community health centers, and sports facilities such as stadiums.

Results: The studied cohort was composed of 187 women and 64 men, with no significant differences in age and BMI between genders. Frailty prevalence varied significantly across clusters, with Cluster 3 showing the highest prevalence (14/47, *P*=.01). Frail individuals resided significantly closer to emergency health centers (960 [SE 904] m vs 1352 [SE 936] m, *P*=.04), main squares/parks (1550 [SE 130] m vs. 2048 [SE 105] m, *P*=.03), and sports fields (3040 [SE 236] m vs 4457 [SE 322]m, *P*=.04) compared with nonfrail individuals. There were no significant differences in urban quality index across frailty groups, but frail individuals lived in areas with higher population density (0.013 [SE 0.001] vs 0.01 [SE 0.0007], *P*=.03).

Conclusions: Frail individuals exhibit geospatial patterns suggesting intentional proximity to health facilities, sports venues, and urban facilities, revealing associations with adaptive responses to frailty and socioeconomic factors. This highlights the crucial intersection of urban environments and frailty, which is important for geriatric medicine and public health initiatives.

(JMIR Aging 2025;8:e64254) doi:10.2196/64254

KEYWORDS

aging; frailty; geospatial clustering; urban factors; neighborhood conditions.

Introduction

Understanding the aging process and the sociodemographic determinants related to enhancing the quality of life has emerged as a very relevant research area in light of the rapid aging of the global population [1-3]. Currently, 12% of the world's population is aged ≥60 years, and projections suggest that this proportion may rise to 21.5% by the mid century [4]. Similarly, the ≥80 years age group is expected to increase from 1.7% to 4.5% [4]. In this context, Chile is experiencing a pronounced aging phenomenon [4,5]. Projections indicate that the Chilean population aged ≥60 years is set to surge from 15.7% to 32.9% by 2050, with the proportion of individuals aged ≥80 years potentially reaching 10.3% [4].

According to the World Health Organization, the frailty syndrome is a crucial determinant regarding the state of dependency, the presence of chronic diseases, polypharmacy, and the quality of life in older people [6,7]. The frailty syndrome is defined as a preventable and reversible clinical state in which the capacity of older people to cope with everyday stressors is compromised by an increase in vulnerability and the physiological deterioration associated with aging [8]. Recent results show a prevalence of frailty in Chile slightly higher than 20% [9,10]. Frail persons have higher risks of mortality, cognitive impairment, fractures, and hospitalization, among other adverse health events, which, considering the increase in the population of older people, represents a challenge for public health and social welfare systems [11,12].

The built environment refers to spaces altered or created by human activities, encompassing a spectrum from homes and schools to workplaces, highways, urban expanses, accessibility to amenities, recreational areas, and pollution [13]. This environment can be delineated into 2 primary components: the microenvironment, encapsulating neighborhood and street-level attributes, and the macroenvironment, which includes the degree of urbanization and patterns of land use [14]. Enhancing our understanding of how the urban physical environment impacts older adults can significantly aid in formulating effective plans and interventions to prevent the progression and onset of frailty while promoting the well-being of this population [15]. Previous research conducted by our group has demonstrated that leveraging geomatic tools, which integrate geolocation as an additional dimension of analysis, can provide valuable insights for studying frailty as a syndrome and supporting the implementation of preventive measures [16,17].

According to reports from the World Health Organization on aging and friendly cities, enhancing the environment through improvements in physical structures and community support is an effective approach to maintaining the health of older people [18]. Recent evidence underscores the impact of neighborhood characteristics on frailty among older people. Those residing in neighborhoods with abundant green spaces exhibit a lower incidence of frailty, whereas individuals perceiving precarious conditions in their surroundings, houses, and environment face a higher risk of frailty [19-21]. A comprehensive multilevel (individual and community) cross-sectional analysis highlighted that older adults living in aesthetically pleasing and walkable neighborhoods tend to exhibit lower levels of frailty. In contrast, areas with high-traffic roads, for example, were associated with a higher prevalence of frailty [22,23]. These findings emphasize the critical role of physical environmental factors in shaping the health and well-being of older populations, highlighting the importance of designing age-friendly communities that promote active and healthy aging.

In this context, this study aims to analyze the relationship between various elements of the urban physical environment and the level of frailty syndrome in older Chilean people.

Methods

Participants and Study Design

The research adopted a cross-sectional case-control design, with a representative sample of older persons (aged ≥65 years old, both men and women) randomly selected from various Family Health Centers and community groups of older people in Talca City, Chile (n=251). All medical centers that serve older adults in the city were considered, ensuring geographical representation. The inclusion criterion was adults aged 65 years and older. Participants with self-reported or medically documented cancer, Parkinson disease, or vascular events were excluded, as were older individuals unable to walk or speak [9]. The calculation of the sample size (aged ≥65 years old, both men and women) considered a prevalence of frailty in older adults of 24.6% [9], with a 95% CI, statistical power of 80%, and a loss percentage of 20%. The proportions of women and men in the sample were determined by the relative distribution of the adult population over 65 years using data from the National Socioeconomic Characterization Survey [24]. No additional stratification was applied.

Frailty Diagnosis

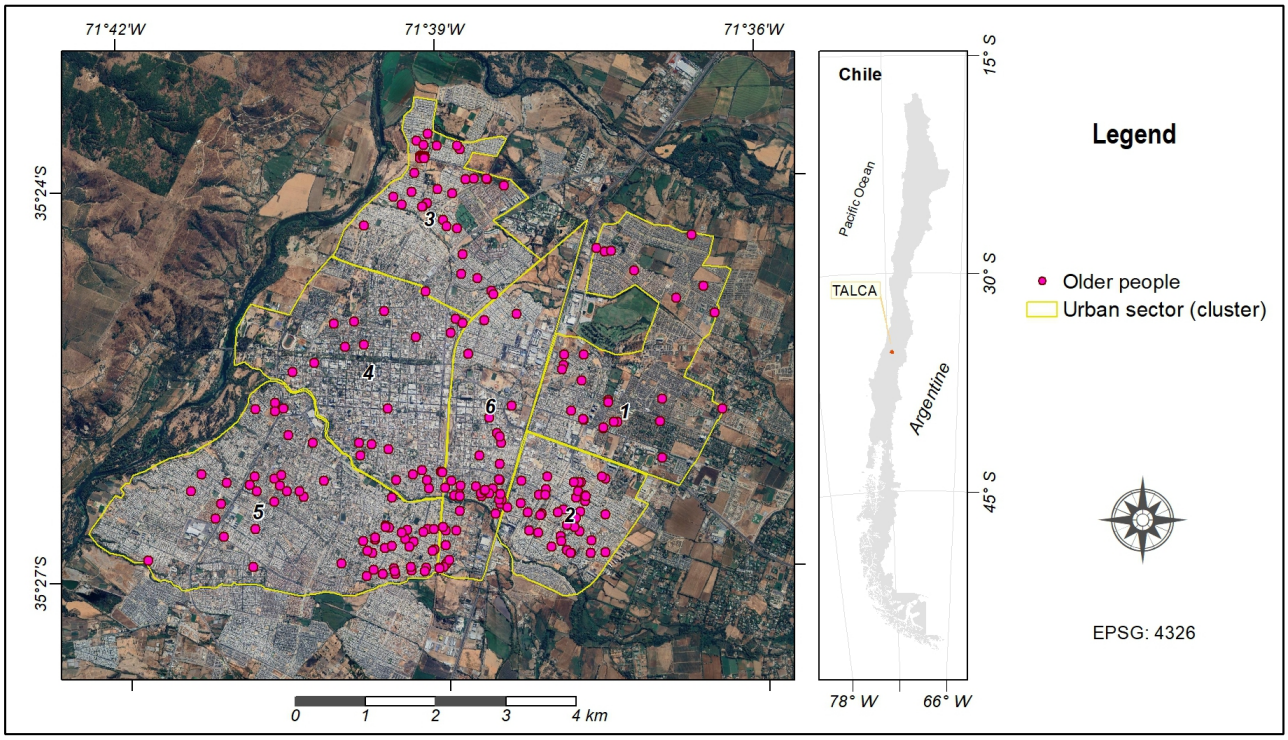
The Fried frailty phenotype criteria were used as the diagnostic tool for assessing frailty [9,25]. This method evaluates the presence or absence of the following components: slowness, weakness, weight loss, exhaustion, and low physical activity. These parameters were defined based on the criteria described previously by Palomo et al [9], which include: (1) slowness: walking velocity below a cut-off of 0.8 m/s average 3-meter walk at a usual pace, adjusted for sex and height according to the standards of the Short Physical Performance Battery, (2) weakness: handgrip strength measured using an Electronic Handgrip Dynamometer (Camry), with sex-specific cut-off (male <27 kg, female <15 kg), (3) weight loss: defined as loss of at least 5 kg in the previous 6 months, (4) exhaustion: a positive response to any of the following two questions from the Center for Epidemiological Studies Depression Scale: “I felt that anything I did was a big effort” and “I felt that I could not keep on doing things” at least 3 to 4 days a week,” (5) low physical activity: difficulty walking, assessed by the questions “Do you have difficulty walking a block?” or "Do you have difficulty climbing several flights of stairs without resting? Participants meeting 3 or more of these components were categorized as frail, those with no 1 or 2 components were considered prefrail, and individuals lacking all components were classified as non-frail or robust [9].

Geospatial Clusters

Each participant was geographically located according to the city address informed and represented as a point object based on the residence information provided during the medical evaluation. All data were organized into a point feature layer accompanied by its corresponding thematic table. This layer was integrated into a geodatabase for subsequent analysis within its geographical context, along with pertinent factors related to the urban physical environment. Figure 1 displays the individual residency positions of each older adult participating in the study within their respective geographical cluster. These geographical

areas were delimited in the city of Talca, Maule Region, Chile, based on geospatial location and sociodemographic characteristics. Cluster 1 corresponds to the northeastern sector characterized by high socioeconomic status. Cluster 2 encompasses the southeastern sector with low socioeconomic status. Cluster 3 covers the northern sector with a lower-middle socioeconomic class, and Cluster 4 represents the "historic center" area. The southern sector of medium-high socioeconomic level is covered by Cluster 5, and Cluster 6 corresponds to the "industrial center" area.

Figure 1. Location of older individuals within the 6 urban sectors.



Urban Quality Level

A georeferenced database was constructed using Geographic Information Systems (GIS) technology to represent pertinent geographical information concerning urban physical facilities within the city, encompassing: (1) vegetables and fruits shops, (2) senior centers or communities, (3) pharmacies, (4) emergency health centers, (5) main squares and parks, (6) family or community health centers, and (7) stadiums and sports fields. Each component within the study area was depicted as a GIS layer, either in point or polygon form, at the neighborhood scale. This representation (Figure 2) was derived from data sourced from OpenStreetMap [26], Google Maps [27] and Infraestructura de Datos Geospaciales de Chile (IDE Chile) [28]. Subsequently, each GIS layer underwent analysis using the Euclidean distance method, providing insights into the proximity of every location within the city to the considered infrastructure. The resulting distance layers were subsequently classified to delineate 3 distinct zones encircling the urban facilities, categorizing their proximity as either close, medium, or distant. Next, each proximity class for every layer was assessed on a scale ranging from 1 to 3, wherein the closest proximity received a score of 3, and the more distant areas were assigned a score of 1. The distance ranges, and the corresponding values assigned to each urban facility were defined according to local context and are presented in Table 1. A general criterion for evaluation was that the closer the facility, the higher the value assigned. A raster calculator was used to aggregate all layers, yielding a summary index where a higher numerical value signifies enhanced urban quality in the depicted area. Afterward, the values derived from the distance analyses and the corresponding summary index for each participant were integrated into the point feature layer. Management, processing, and analyzing data were performed using ArcGIS software, version 10 (ESRI, Redlands, USA).

Figure 2. Geographical distribution of urban physical facilities within Talca City.

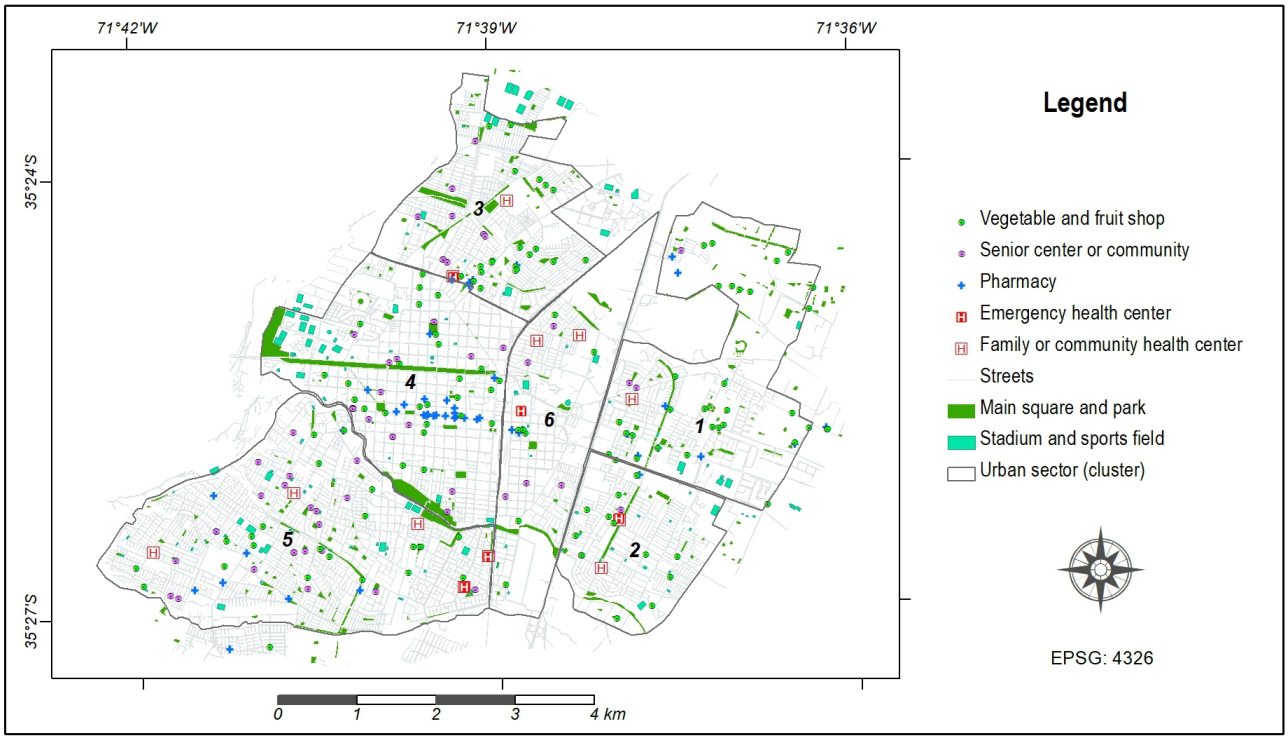


Table . Distance ranges and values for each urban physical facility.

Facility and distance ranges	Value
Vegetables and fruit shops (m)	
<300	3
300 - 600	2
>600	1
Senior centers or communities (m)	
<500	3
500 - 1000	2
>1000	1
Pharmacies (m)	
<500	3
500 - 1000	2
>1000	1
Emergency health centers (m)	
<1000	3
1000 - 2000	2
>2000	1
Main squares and parks (m)	
<200	3
200 - 400	2
>400	1
Family or community health centers (m)	
<700	3
700 - 1400	2
>1400	1
Stadiums and sports fields (m)	
<400	3
400 - 600	2
>600	1

Statistical Analysis

Statistical analyses were conducted using GraphPad Prism 9. Continuous variables were expressed as mean (SD) or median (95% CI). Categorical variables were expressed as percentages with a 95% CI. In the evaluation of differences between groups, the chi-square test with Yate correction was used to assess proportions, while ANOVA or the Kruskall-Wallis test, as appropriate, was applied to assess differences in means or medians. Statistical significance was considered at *P* values below .05.

Ethical Considerations

The institutional board review approval for this study was obtained from the Comité de Ética Científica (CEC) of Universidad de Talca (reference number 06 - 2021). All procedures followed adhered to the ethical standards of the CEC and the World Medical Association's Declaration of Helsinki. All study participants provided written informed consent.

Results

Sociodemographic Characteristics and Cluster Distribution

Table 2 presents the sociodemographic characteristics of the analyzed cohort of older people. The sample comprised 74.5% women and 25.2% men, with no significant differences observed in age and BMI between the 2 genders. In addition, Table 2 illustrates the distribution of the geospatial clusters established during the study. Analysis indicated no significant difference in the distribution of the 6 designated clusters between men and women.

Table . Sociodemographic description and geospatial distribution of the studied sample of older people.

Variable	Women (n=187)	Men (n=64)	*P* value
Gender, % (95% CI)	74.5 (68.7-79.4)	25.5 (20.5-31.2)	—[a]
Age (years), mean (SD)	73.8 (5.9)	75 (5)	.152
BMI (kg/m^2), mean (SD)	33.6 (33.5)	28.4 (6.5)	.243
Spatial cluster showing percentage of prevalence, % (95% CI)			
Cluster 1	10.2 (6.6-15.3)	15.6 (8.7-26.4)	.260
Cluster 2	18.7 (13.7-24.9)	14.1 (7.6-24.6)	.452
Cluster 3	16.6 (11.9-22.5)	25 (16-36.8)	.142
Cluster 4	11.2 (7.5-16.6)	9.4 (4.4-18.9)	.817
Cluster 5	26.2 (20.4-32.9)	17.2 (9.8-28.2)	.176
Cluster 6	17.1 (12.4-23.2)	18.8 (11.1-29.9)	.845

[a] —: not applicable.

Analysis of Urban Quality

Figure 3 shows the different levels of urban quality [3] obtained from the cumulative assessment of physical environmental elements considered in this study. These elements mainly encompass basic urban services and infrastructures essential for the local population, with particular importance for the cohort of older people under study. The quality level is closely related to the accessibility of the various facilities from each location within the city. In addition, Figure 3 also shows the individual distribution of older adults, each denoted by their frailty status, which is subsequently analyzed in Table 3. This table displays the distribution of the frailty status through the different geospatial clusters. The prevalence of frailty varies between 7.3% and 34.1% among these clusters. Notably, cluster 3 exhibits a significantly high prevalence of frail people (34.1%, *P*=.006), followed by cluster 5 (21.9%, *P*=.417); however, this last prevalence is not significantly high.

Figure 3. Frailty status of older individuals over summary index within the geospatial clusters defined for Talca city.

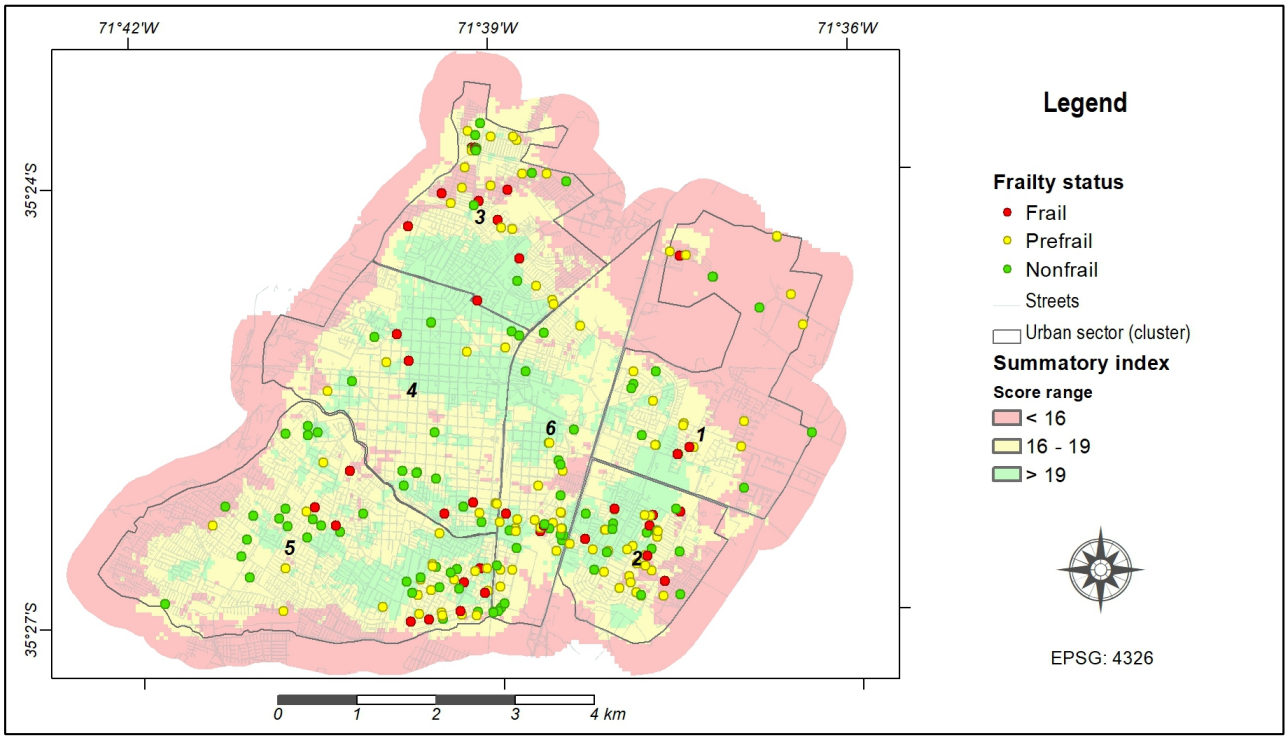


Table . Distribution of older people by frailty status and geospatial clustering.

Spatial cluster	Frailty status considering the percentage of prevalence			*P* value
	Nonfrail (%, 95% CI)	Prefrail (%, 95% CI)	Frail (%, 95% CI)	
Cluster 1	14.3 (8.5 - 22.9)	11.7 (6.9 - 19)	7.3 (2.5 - 19.4)	.620
Cluster 2	14.3 (8.5 - 22.9)	21.6 (14.9 - 30.2)	17.7 (8.5 - 31.3)	.270
Cluster 3	12.1 (6.8 - 20.4)	19.8 (13.5 - 28.2)	34.1 (21.5 - 49.5)	.006
Cluster 4	14.3 (8.5 - 22.9)	8.1 (4.3 - 14.6)	12.2 (5.3 - 25.5)	.477
Cluster 5	30.7 (22.2 - 40.9)	20.7 (14.2 - 29.2)	21.9 (12 - 36.7)	.417
Cluster 6	14.3 (8.5 - 22.9)	18 (11.9 - 26.2)	7.3 (2.5 - 19.4)	.225
Total	100	100	100	—[a]

[a] —: not applicable.

Analysis of Distances to Urban Facilities, Summary Index, and Population Density

Figure 4 illustrates the variation in average distance between old persons (categorized by their frailty status) and relevant urban facilities. For the facilities of vegetable and fruit shops (A), senior centers or communities (B), pharmacies (C), and family and community health centers (F), there were no significant differences in the average distance across different frailty status groups. However, a clear linear trend is observed between the groups, where frail people tend to reside further from vegetable and fruit shops (frail: 335.6 [SE 31.2] vs nonfrail: 275.9 [SE 16.5]) and closer to the senior centers or communities (frail: 368.2 [SE 38.6] vs nonfrail: 435.9 [SE 35.9]) than the nonfrail people. On the other hand, the facilities of emergency health centers (D), main squares and parks (E), and stadiums and sports fields (G) present significant differences in the average distance across different frailty status groups, where frail people resided significantly closer to emergency health centers (frail: 960.4 [SE 90.4] vs nonfrail: 1352 [SE 93.6], *P*=.04), main squares and parks (frail: 155 [SE 13] vs prefrail: 204.8 [SE 10.5], *P*=.03), and stadiums and sports fields (frail: 304 [SE 23.6] vs prefrail: 445.7 [SE 32.2], *P*=.04), than both nonfrail and prefrail people, respectively. Finally, Figure 5 shows the variations in the summary index and population density between the frailty status groups. While the summary index (A) shows no significant differences between groups, there is a significant difference in population density (B) between the frail and nonfrail status groups, being higher for frail people (frail: 0.013 [SE 0.001] vs nonfrail: 0.01 [SE 0.0007], *P*=.03).

Figure 4. Comparison of the mean distance for different frailty status diagnosed using the Fried phenotype criteria with respect to relevant urban facilities of (A) vegetables and fruits shops; (B) senior centers or communities; (C) pharmacies; (D) emergency health centers; (E) main squares and parks; (F) family or community health centers; (G) stadiums and sports fields. The data presented are the mean (SE). Statistical analysis was performed using the ANOVA test with the Tukey pos-hoc test. *P<.05.

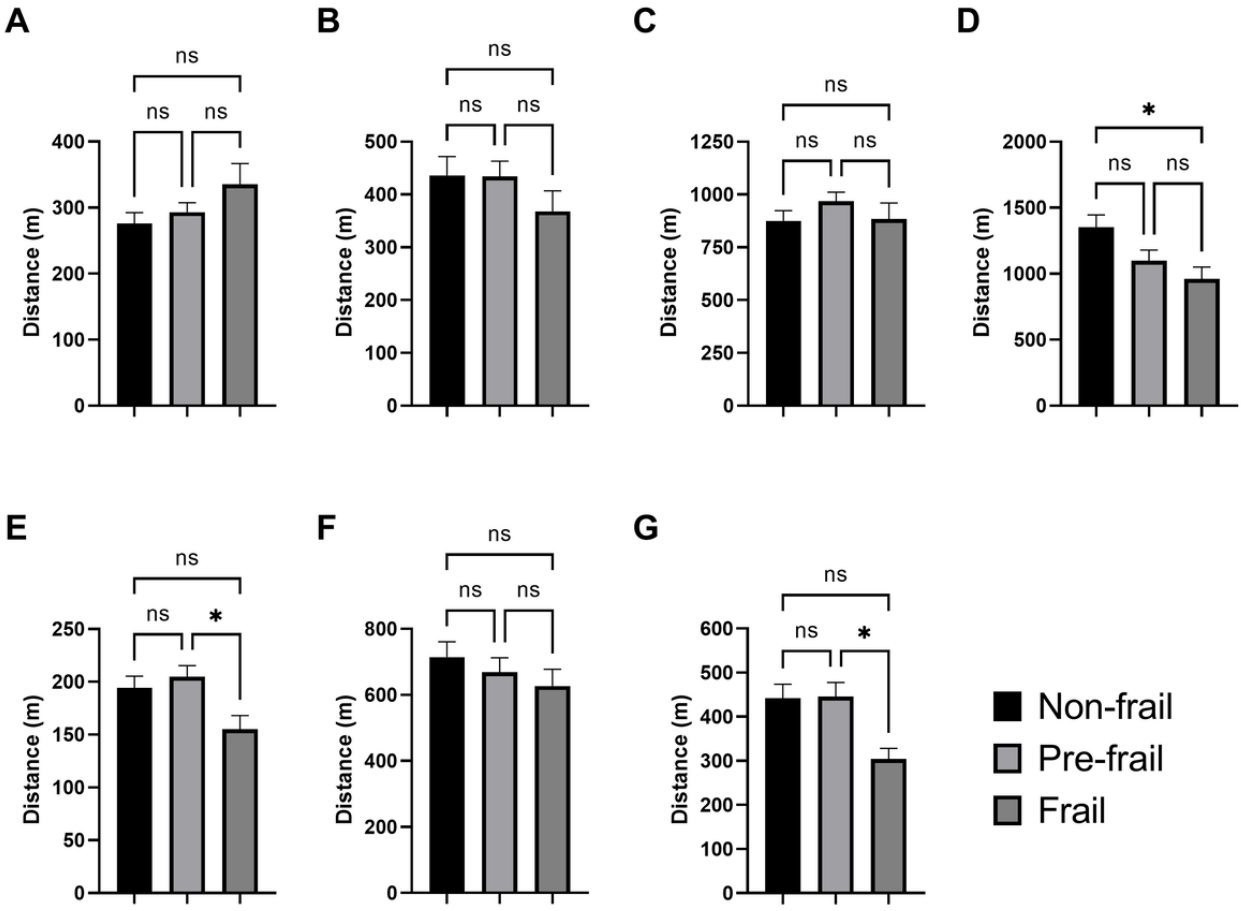
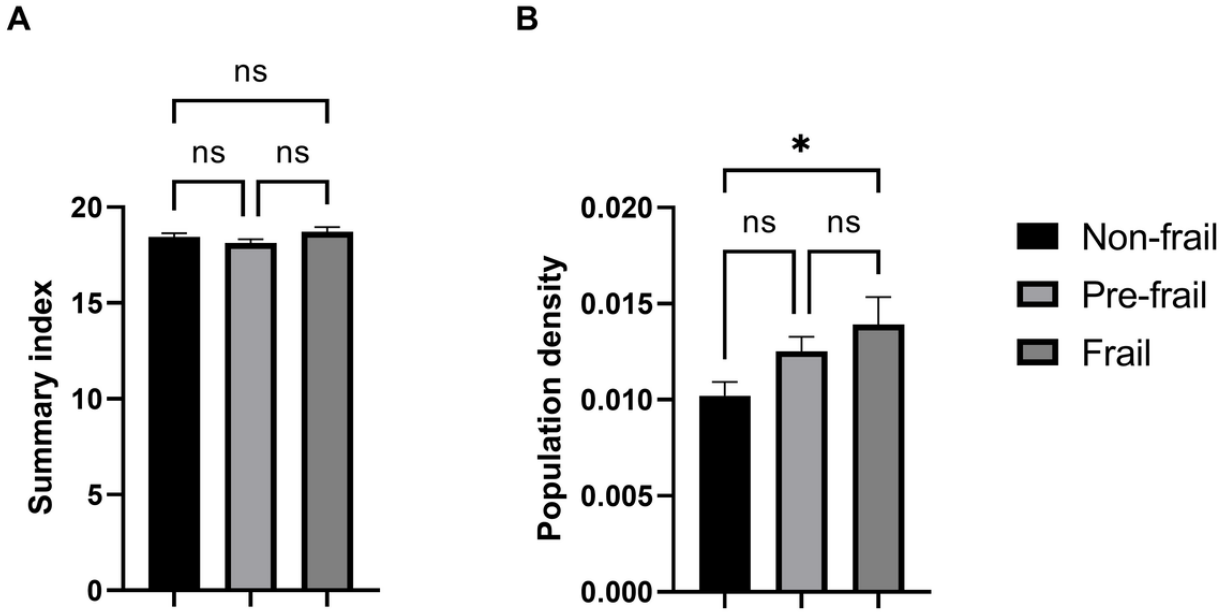


Figure 5. Comparison of mean values for different frailty status diagnosed according to the Fried phenotype criteria for the (A) summary index, (B) population density. The data presented are mean (SE). Statistical analysis was performed using the ANOVA test with the Tukey post-hoc test. *P<.05.



Discussion

Principal Findings

This study explores the relationship between urban physical environment factors and frailty syndrome, with a particular focus on its diagnostic criteria. The results mentioned above indicate that frail individuals tend to live closer to emergency health centers, stadiums, sports fields, and senior centers, suggesting efforts to improve their health and social engagement. In addition, frail individuals primarily reside in densely populated areas, which are associated with limited physical activity and higher mortality rates. Given the relevance of the Fried frailty phenotype as a standard frailty assessment tool in various studies involving older adults, it was selected for frailty characterization [11,29,30]. Investigating the link between frailty and urban environment is innovative and can provide invaluable insights for government programs aimed at enhancing the well-being of older people [31,32]. However, this emerging topic remains underdeveloped, limiting opportunities for comparisons specific to frailty. The incorporation of geomatics tools enables the use of geolocation as an additional analytical dimension, offering valuable insights into the study of frailty as a syndrome. This is especially important when the primary focus is on understanding the impact of the urban physical environment on the frail condition of older individuals [16,17]. Currently, the use of information and communication technologies among older persons is increasing, and different benefits have been developed about frailty with respect to predicting the risk, assisting in identifying changes in frailty parameters, enhancing adherence to a healthy diet, and distinguishing older rehabilitation patients who need to be readmitted to the hospital from those who can remain in the community which could facilitate this key process [33-37].

The data from Table 2 indicate homogeneity within the cohort, with no significant influences in the study parameters when stratified by gender. Participants had an average age range of 73 - 75 years old and exhibited a high BMI. This trend of elevated BMI is consistent with findings from other studies focused on older Chilean people, wherein a significant prevalence of obesity (BMI >25 kg/m^2) has been documented [8,9,38]. Gender did not appear to influence the distribution prevalence within the identified clusters. However, it is necessary to account for gender in geospatial clustering analyses involving older people, given its potential impact on various health outcomes, as recently evidenced in COVID-19 studies [39].

The findings presented in Table 3 suggest that geospatial distribution significantly impacts the prevalence of frailty syndrome. Notably, clusters 3 and 5, which are associated with middle to low socioeconomic classes, concentrate more than half of the frail people. This evidence is consistent with studies where low socioeconomic groups have been associated with a higher risk of developing frailty [31]. Furthermore, previous research indicates that frailty condition among older people tends to exhibit spatial clustering, wherein certain areas within and outside the city display localized concentrations of both high and low prevalence of frail people [16,40]. These clustering patterns can be attributed to the diverse urban infrastructures and socioeconomic disparities observed across different sectors within the city [16].

In this study, the outcomes depicted in Figure 3 reveal a geospatial clustering concentrated in the southern part of the city, primarily associated with clusters 3 and 5. However, when considering the general condition of the urban physical environment, characterized by the summary index, the group of frail people does not show significant differences compared with the other groups (Figure 5A). This fact suggests that, despite various groups of adults residing in environments characterized by comparable habitability conditions, distinctions in certain factors, as evidenced in Figure 4, can exert relevant positive or negative influences. Thus, the urban physical environment is likely connected with concurrent factors associated with the frailty syndrome, serving either as causal factors or as responses to such influences.

Frailty is acknowledged as a critical factor influencing the health and well-being of older people [11-18]. It stands out as the primary risk factor and indicator for the initiation of dependency, as well as the occurrence of chronic diseases, hospitalizations, falls, fractures, and mortality. The extensive array of health-related challenges observed in frail people may be associated with their proximity to emergency health centers, as well as family or community health centers [15-20]. Likewise, frailty has exhibited strong associations with obesity, sarcopenia, and low physical activity. Consequently, the proximity identified between the frail group and stadiums and sports fields may suggest a purposeful endeavor to improve body composition and address underlying clinical conditions [41,42].

Simultaneously, the frail group exhibited greater proximity to main squares and parks and to senior centers or communities renowned for fostering social activities among older people [43]. Given the well-established associations between frailty and social isolation, depression, and loneliness [44], the observed significant trend might signify a proactive response to enhance community engagement. This trend could also be indicative of the city government's concerted efforts to provide these facilities to a population in need. By strategically situating frail individuals near these spaces, urban planners and policymakers may be aiming to reduce the adverse effects of isolation and encourage greater social participation. Engaging in recreational activities and interacting with peers in accessible public spaces can help mitigate the mental health challenges often faced by frail older adults. Furthermore, these spaces serve as venues for both physical and mental stimulation, which are critical to maintaining functional independence and improving overall well-being.

Conversely, the robust association between frailty and obesity might elucidate the trend observed trend of increased distance from the frail group to fruit and vegetable shops. Nevertheless, validation of this hypothesis necessitates examination within a more extensive cohort. While elevated intake of fruits and vegetables has been linked to a diminished frailty risk, the impact of the proximity of these food supply points on frailty remains ambiguous [45]. Furthermore, greater distances to pharmacies may indicate challenges in accessing essential drugs

and medications, a factor that should be taken into account in health programs tailored to support the frailty group.

The findings depicted in Figure 5B indicate that the frail group tends to inhabit regions characterized by elevated population density. Older adults tend to live in densely populated marginalized areas due to economic, social, and structural factors. Financial limitations, the lack of adequate housing options, access to family support networks, and mobility barriers are some of the key factors influencing their stay in these areas. This circumstance poses an increased risk to this group, as current scientific understanding suggests a positive correlation between population density in a given neighborhood and increased susceptibility of middle-aged and older adults to overweight conditions [46]. This association can be ascribed to the proclivity of individuals residing in densely populated areas to adopt a sedentary lifestyle, marked by limited physical activity and diminished energy expenditure [20,46]. Notably, low physical activity and sedentarism are integral components of the frailty phenotype observed across diverse cohorts of older adults [11,12,26]. Furthermore, elevated population density has been lined to increased mortality rates across all causes in older people [11,12,47,48].

Initiatives aimed at preventing frailty underscore the imperative to advocate for physical activity, nutrition, and social engagement as primary and efficacious interventions. These interventions can be effectively implemented through a health education program tailored to inspire and engage older people [49,50]. Our findings underscore the importance of taking urban factors into account when examining the frailty condition of older people. This evidence aligns with previous research emphasizing the importance of these factors in the well-being and social engagement of older adults, as well as their association with frailty [16,51]. These results emphasize the imperative to investigate further and enhance our comprehension of the role played by urban factors in shaping frailty among older people. Regardless, it is imperative to consider the complex interrelationship between the urban physical environment and frailty condition when devising structural preventive measures aimed at improving the well-being of older people [16,47].

Conclusions

Contemporary evidence underscores the relevance of urban factors in influencing the onset of frailty and the diverse health factors linked to this syndrome. Frailty stands as a highly prevalent geriatric syndrome in the elderly, elevating the susceptibility to a range of adverse health and social outcomes. Addressing this challenge necessitates the development of age-friendly cities tailored to the needs of older populations. Our findings suggest that individuals classified as frail tend to reside in closer proximity to emergency health centers, as well as family or community health centers, which may be indicative of adaptive responses to the features associated with frailty, such as elevated mortality risk and diminished levels of physical activity. Likewise, frail people tend to reside closer to stadiums and sports fields, which may imply a deliberate endeavor to enhance body composition and address underlying clinical conditions. On its part, the closer proximity of frail people to urban infrastructures such as main squares and parks, and senior centers or communities may be indicative of the concerted effort by the municipal government to provide these facilities to a population in need. However, it is important to highlight that the frail group predominantly inhabits areas characterized by elevated population density, aligning with earlier research associating higher mortality rates with increased population density among older adults. Likewise, the geospatial clustering highlights the relevance of both socioeconomic status and geographic location in relation to frailty prevalence, unveiling an elevated occurrence of this syndrome in sectors characterized by a lower-middle socioeconomic class. This finding aligns with and reinforces previously established evidence. Ultimately, the existing body of evidence, coupled with our study findings, underscores the significance of investigating frailty and its associations with the urban environment and related factors. This emerging field of research holds groundbreaking potential, offering substantial implications for geriatric medicine. Furthermore, it provides invaluable insights that can inform the development of governmental initiatives aimed at promoting healthy aging and proactively preventing frailty.

Acknowledgments

This work has been funded by Interuniversity Center for Healthy Aging, (Code RED211993). Also, the authors would like to thank ANID-FONDECYT 1211136 and ANID-FONDECYT 3230783.

Authors' Contributions

DA, YO, and JCC contributed to conceptualization. DA and YO handled data curation and formal analysis. CM, IP, and EF managed funding acquisition. DA, YO, and JCC conducted investigation. DA, YO, and JCC managed methodology. CM, IP, and EF handled project administration. CM and IP managed resources. YO handled software. CM and IP supervised. CM and IP managed validation. DA and YO conducted writing-original draft and writing-review & editing.

Conflicts of Interest

None declared.

References

1. He W, Goodkind D, Kowal PR. An aging world: 2015. : United States Census Bureau Washington, DC; 2016.

2. Nguyen TN, Cumming RG, Hilmer SN. A review of frailty in developing countries. J Nutr Health Aging 2015 Nov;19(9):941-946. [doi: 10.1007/s12603-015-0503-2] [Medline: 26482697]
3. Bloom DE, Boersch-Supan A, McGee P, Seike A. Population aging: facts, challenges, and responses. Benefits and compensation international 2011;41(1):22.
4. Thumala D, Kennedy BK, Calvo E, et al. Aging and health policies in Chile: new agendas for research. Health Syst Reform 2017 Oct 2;3(4):253-260. [doi: 10.1080/23288604.2017.1353844] [Medline: 30359182]
5. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139(6-11):6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
6. WHO guidelines on integrated care for older people (ICOPE). World Health Organization. 2018. URL: https://www.who.int/ageing/publications/guidelines-icope/en [accessed 2025-04-14]
7. Arauna D, Cerda A, García-García JF, et al. Polypharmacy Is Associated with Frailty, Nutritional Risk and Chronic Disease in Chilean Older Adults: Remarks from PIEI-ES Study. Clin Interv Aging 2020;15:1013-1022. [doi: 10.2147/CIA.S247444] [Medline: 32636616]
8. WHO. WHO clinical consortium on healthy ageing: topic focus: frailty and intrinsic capacity. 2017 Presented at: Report of Consortium Meeting; Dec 1-2, 2016; Geneva, Switzerland URL: https://www.who.int/publications/i/item/WHO-FWC-ALC-17.2 [accessed 2024-12-10]
9. Palomo I, Giacaman RA, León S, et al. Analysis of the characteristics and components for the frailty syndrome in older adults from central Chile. The PIEI-ES study. Arch Gerontol Geriatr 2019;80(70-5):70-75. [doi: 10.1016/j.archger.2018.10.004] [Medline: 30390427]
10. Palomo I, García F, Albala C, et al. Characterization by gender of frailty syndrome in elderly people according to FrailTrait Scale and Fried Frailty Phenotype. J Pers Med 2022 Apr 29;12(5):712. [doi: 10.3390/jpm12050712] [Medline: 35629135]
11. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
12. García-García FJ, Carcaillon L, Fernandez-Tresguerres J, et al. A new operational definition of frailty: the Frailty Trait Scale. J Am Med Dir Assoc 2014 May;15(5):371. [doi: 10.1016/j.jamda.2014.01.004] [Medline: 24598478]
13. Pasala SK, Rao AA, Sridhar GR. Built environment and diabetes. Int J Diabetes Dev Ctries 2010 Apr;30(2):63-68. [doi: 10.4103/0973-3930.62594] [Medline: 20535308]
14. Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. Prev Med 1999 Dec;29(6 Pt 1):563-570. [doi: 10.1006/pmed.1999.0585] [Medline: 10600438]
15. Di Ciaula A, Portincasa P. The environment as a determinant of successful aging or frailty. Mech Ageing Dev 2020 Jun;188(111244):111244. [doi: 10.1016/j.mad.2020.111244] [Medline: 32335099]
16. Mena C, Ormazabal Y, Fuentes E, Palomo I. Impacts of physical environment perception on the frailty condition in older people. Geospat Health 2020 Jun 18;15(1):32575969. [doi: 10.4081/gh.2020.888] [Medline: 32575969]
17. Mena C, Fuentes E, Ormazábal Y, Palomo-Vélez G, Palomo I. Effect of straight-line and road network distances to parks and markets on anthropometric measurements, biochemical markers, and a healthy lifestyle in adult people. Sport Sci Health 2016 Apr;12(1):55-61. [doi: 10.1007/s11332-015-0253-x]
18. World Health Organization. Decade of healthy ageing: baseline report. 2020 URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2024-12-10]
19. Gobbens RJJ. Cross-sectional and longitudinal associations of environmental factors with frailty and disability in older people. Arch Gerontol Geriatr 2019;85:103901. [doi: 10.1016/j.archger.2019.103901] [Medline: 31352186]
20. Yu R, Wang D, Leung J, Lau K, Kwok T, Woo J. Is neighborhood green space associated with less frailty? Evidence from the Mr. and Ms. Os (Hong Kong) study. J Am Med Dir Assoc 2018 Jun;19(6):528-534. [doi: 10.1016/j.jamda.2017.12.015] [Medline: 29402649]
21. Duppen D, Van der Elst MCJ, Dury S, Lambotte D, De Donder L, D-SCOPE. The social environment's relationship with frailty: evidence from existing studies. J Appl Gerontol 2019 Jan;38(1):3-26. [doi: 10.1177/0733464816688310] [Medline: 28380715]
22. Abe T, Carver A, Sugiyama T. Associations of neighborhood built and social environments with frailty among mid-to-older aged Australian adults. Geriatr Gerontol Int 2021 Oct;21(10):893-899. [doi: 10.1111/ggi.14253] [Medline: 34355479]
23. Ye B, Gao J, Fu H. Associations between lifestyle, physical and social environments and frailty among Chinese older people: a multilevel analysis. BMC Geriatr 2018 Dec 14;18(1):314. [doi: 10.1186/s12877-018-0982-1] [Medline: 30547760]
24. CASEN 2017. Base de datos Casen 2017 SPSS (Base de datos) [Website in Spanish]. URL: https://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2017 [accessed 2025-04-14]
25. Albala C, Lera L, Sanchez H, et al. Frequency of frailty and its association with cognitive status and survival in older Chileans. Clin Interv Aging 2017;12(995-1001):995-1001. [doi: 10.2147/CIA.S136906] [Medline: 28721027]
26. Open street map. OSM Foundation. URL: https://www.openstreetmap.org [accessed 2024-11-01]
27. Google Maps Google LLC. URL: https://www.google.com/maps [accessed 2024-11-01]
28. Geospatial infrastructure chile. IDE Chile. URL: https://www.ide.cl [accessed 2024-11-01]

29. Martin FC, O'Halloran AM. Tools for assessing frailty in older people: general concepts. Adv Exp Med Biol 2020;1216(9-19):9-19. [doi: 10.1007/978-3-030-33330-0_2] [Medline: 31894542]
30. Morley JE, Vellas B, van Kan GA, et al. Frailty consensus: a call to action. J Am Med Dir Assoc 2013 Jun;14(6):392-397. [doi: 10.1016/j.jamda.2013.03.022] [Medline: 23764209]
31. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
32. Dent E, Martin FC, Bergman H, Woo J, Romero-Ortuno R, Walston JD. Management of frailty: opportunities, challenges, and future directions. Lancet 2019 Oct 12;394(10206):1376-1386. [doi: 10.1016/S0140-6736(19)31785-4] [Medline: 31609229]
33. Kim B, McKay SM, Lee J. Consumer-grade wearable device for predicting frailty in Canadian home care service clients: prospective observational proof-of-concept study. J Med Internet Res 2020 Sep 3;22(9):e19732. [doi: 10.2196/19732] [Medline: 32880582]
34. Ramezani R, Zhang W, Xie Z, et al. A combination of indoor localization and wearable sensor-based physical activity recognition to assess older patients undergoing subacute rehabilitation: baseline study results. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e14090. [doi: 10.2196/14090] [Medline: 31293244]
35. Valdés-Aragonés M, Pérez-Rodríguez R, Carnicero JA, et al. Effects of monitoring frailty through a mobile/web-based application and a sensor kit to prevent functional decline in frail and prefrail older adults: FACET (frailty care and well function) pilot randomized controlled trial. J Med Internet Res 2024 Oct 22;26:e58312. [doi: 10.2196/58312] [Medline: 39436684]
36. Su Y, Wu KC, Chien SY, Naik A, Zaslavsky O. A mobile intervention designed specifically for older adults with frailty to support healthy eating: pilot randomized controlled trial. JMIR Form Res 2023 Nov 15;7:e50870. [doi: 10.2196/50870] [Medline: 37966877]
37. Keränen NS, Kangas M, Immonen M, et al. Use of information and communication technologies among older people with and without frailty: a population-based survey. J Med Internet Res 2017 Feb 14;19(2):e29. [doi: 10.2196/jmir.5507] [Medline: 28196791]
38. Bravo C, Godoy J, Sánchez Y, Riveros MJ. Asociación entre el nivel de autovalencia y el estado nutricional en adultos mayores chilenos [Article in Spanish]. Rev chil nutr 2021;48(5):741-747. [doi: 10.4067/S0717-75182021000500741]
39. Ramírez-Aldana R, Gomez-Verjan JC, Bello-Chavolla OY, García-Peña C. Spatial epidemiological study of the distribution, clustering, and risk factors associated with early COVID-19 mortality in Mexico. PLoS One 2021;16(7):e0254884. [doi: 10.1371/journal.pone.0254884] [Medline: 34288952]
40. Taylor D, Barrie H, Lange J, Thompson MQ, Theou O, Visvanathan R. Geospatial modelling of the prevalence and changing distribution of frailty in Australia - 2011 to 2027. Exp Gerontol 2019 Aug;123:57-65. [doi: 10.1016/j.exger.2019.05.010] [Medline: 31129145]
41. Crow RS, Lohman MC, Titus AJ, et al. Association of obesity and frailty in older adults: NHANES 1999-2004. J Nutr Health Aging 2019;23(2):138-144. [doi: 10.1007/s12603-018-1138-x] [Medline: 30697622]
42. Yuan L, Chang M, Wang J. Abdominal obesity, body mass index and the risk of frailty in community-dwelling older adults: A systematic review and meta-analysis. Age Ageing 2021 Jun 28;50(4):1118-1128. [doi: 10.1093/ageing/afab039] [Medline: 33693472]
43. Noon RB, Ayalon L. Older adults in public open spaces: age and gender segregation. Gerontologist 2018 Jan 18;58(1):149-158. [doi: 10.1093/geront/gnx047] [Medline: 28535273]
44. Davies K, Maharani A, Chandola T, Todd C, Pendleton N. The longitudinal relationship between loneliness, social isolation, and frailty in older adults in England: a prospective analysis. Lancet Healthy Longev 2021 Feb;2(2):e70-e77. [doi: 10.1016/S2666-7568(20)30038-6] [Medline: 36098160]
45. Kojima G, Avgerinou C, Iliffe S, Jivraj S, Sekiguchi K, Walters K. Fruit and vegetable consumption and frailty: a systematic review. J Nutr Health Aging 2018;22(8):1010-1017. [doi: 10.1007/s12603-018-1069-6] [Medline: 30272107]
46. Wang R, Feng Z, Xue D, Liu Y, Wu R. Exploring the links between population density, lifestyle, and being overweight: secondary data analyses of middle-aged and older Chinese adults. Health Qual Life Outcomes 2019 Jun 11;17(1):100. [doi: 10.1186/s12955-019-1172-3] [Medline: 31186044]
47. Meijer M, Kejs AM, Stock C, Bloomfield K, Ejstrud B, Schlattmann P. Population density, socioeconomic environment and all-cause mortality: a multilevel survival analysis of 2.7 million individuals in Denmark. Health Place 2012 Mar;18(2):391-399. [doi: 10.1016/j.healthplace.2011.12.001] [Medline: 22209409]
48. Verghese J, Ayers E, Sathyan S, et al. Trajectories of frailty in aging: prospective cohort study. PLOS ONE 2021;16(7):e0253976. [doi: 10.1371/journal.pone.0253976] [Medline: 34252094]
49. Shinkai S, Yoshida H, Taniguchi Y, et al. Public health approach to preventing frailty in the community and its effect on healthy aging in Japan. Geriatr Gerontol Int 2016 Mar;16 Suppl 1(87-97):87-97. [doi: 10.1111/ggi.12726] [Medline: 27018287]
50. Cesari M, Prince M, Thiyagarajan JA, et al. Frailty: an emerging public health priority. J Am Med Dir Assoc 2016 Mar 1;17(3):188-192. [doi: 10.1016/j.jamda.2015.12.016] [Medline: 26805753]

51. Dawson-Townsend K. Social participation patterns and their associations with health and well-being for older adults. SSM Popul Health 2019 Aug;8(100424):100424. [doi: 10.1016/j.ssmph.2019.100424] [Medline: 31312713]

Abbreviations

ANOVA: analysis of variance
CEC: Comité de Ética Científica
GIS: geographic information systems
IDE: Infraestructura de Datos Geospaciales

Edited by A Garcia; submitted 23.07.24; peer-reviewed by J Valenzuela, TTT Nguyen; revised version received 02.12.24; accepted 15.12.24; published 17.04.25.

Please cite as:
Ormazábal Y, Arauna D, Cantillana JC, Palomo I, Fuentes E, Mena C
Unveiling the Frailty Spatial Patterns Among Chilean Older Persons by Exploring Sociodemographic and Urbanistic Influences Based on Geographic Information Systems: Cross-Sectional Study
JMIR Aging 2025;8:e64254
URL: https://aging.jmir.org/2025/1/e64254
doi:10.2196/64254

A Comparison of Patient and Provider Perspectives on an Electronic Health Record–Based Discharge Communication Tool: Survey Study

Dorothy Yingxuan Wang[1,2], PhD; Eliza Lai-Yi Wong[1,2], PhD; Annie Wai-Ling Cheung[1,2], MPhil; Kam-Shing Tang[3], MD; Eng-Kiong Yeoh[1,2], MBBS

[1]JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Centre for Health Systems & Policy Research, JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Kwong Wah Hospital, Hospital Authority, Hong Kong, China (Hong Kong)

Corresponding Author:
Eliza Lai-Yi Wong, PhD
JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Abstract

Background: Hospital discharge for older adult patients carries risks. Effective patient-provider communication is crucial for postacute care. Technology-based communication tools are promising in improving patient experience and outcomes. However, there is limited evidence comparing patient and provider user experiences on a large-scale basis, hindering the exploration of true patient-provider shared understanding.

Objective: This study aimed to evaluate an electronic health record–based discharge communication tool by examining and comparing patient and provider perspectives.

Methods: This study comprised a cross-sectional self-administered staff survey and a pre-post cross-sectional patient survey. Physicians, nurses, and older adult patients aged 65 years and older discharged from 4 public hospitals were included. Patient-provider comparison items focused on 3 aspects of the design quality of the tool (information clarity, adequacy, and usefulness) and overall satisfaction with the tool. In addition, patients' experience of discharge information and their medication-taking behaviors before and after the program implementation were compared based on a validated local patient experience survey instrument. Providers' perceived usefulness of this tool to their work and implementation intentions were measured based on the technology acceptance model to enhance understanding of their experiences by conducting structural equation modeling analysis.

Results: A total of 1375 and 2353 valid responses were received from providers and patients, respectively. Patients' overall satisfaction with this communication tool is significantly higher than providers', and patients rated the information clarity and usefulness presented by this tool higher as well (P<.001). However, patients rated information adequacy significantly lower than providers (P<.001). Meanwhile, patients reported a significant improvement in their experience of discharge medication information, and fewer patients reported side effects encounters after the program implementation (126/1083, 11.6% vs 111/1235, 9%; P=.04). However, providers showed inconsistent implementation fidelity. Providers' perceived quality of the tool design (β coefficient=0.24, 95% CI 0.08-0.40) and perceived usefulness to their work (β coefficient=0.57, 95% CI 0.43-0.71) significantly impacted their satisfaction. Satisfaction can significantly impact implementation intentions (β coefficient=0.40, 95% CI 0.17-0.64), which further impacts implementation behaviors (β coefficient=0.16, 95% CI 0.10-0.23).

Conclusions: A notable disparity exists between patients and health care providers. This may hinder the achievement of the tool's benefits. Future research should aim for a comprehensive overview of implementation barriers and corresponding strategies to enhance staff performance and facilitate patient-provider shared understanding.

(JMIR Aging 2025;8:e60506) doi:10.2196/60506

KEYWORDS

older adult; gerontology; geriatric; old; older; elderly; aging; aged; post-acute care; communication; satisfaction; medication information; patient-provider comparison; technology-based intervention; technology acceptance model; discharge; EHR; record; portal; cross-sectional; survey; questionnaire; experience; attitude; opinion; perception; perspective; acceptance; adoption; design; user experience

Introduction

At hospital discharge, health care providers play a crucial role in delivering comprehensive medication information, including side effects and warnings, to ensure safe medication therapy during patients' postacute care [1]. However, previous literature has reported that older adult patients often lack awareness or understanding of their medication regimen after being discharged home [2-4]. Insufficient knowledge is associated with their suboptimal adherence to treatment [5,6], an elevated likelihood of adverse events [7], increased readmissions and emergency department visits [8], and burden on the health care system [9].

A wide array of communication strategies has been documented in the literature with the aim of facilitating information provision by health care providers and enhancing patient awareness and understanding of health-related information [10]. Their effectiveness in reducing readmissions and enhancing patient satisfaction was supported by a recent meta-analysis [1]. Notably, information technology–based communication practices have emerged as a prominent and preferred mode for delivering discharge information, as highlighted in literature reviews [11]. In addition, a systematic review concluded that computer-enabled discharge communication interventions improve both patient and provider satisfaction and reduce perceived adverse events [12]. A Cochrane review further indicated that computer-generated reminders presented on paper can enhance the quality of care [13]. However, there is a scarcity of research comparing the perspectives of older adult patients and health care providers with concordance measures for a large-scale technology-based discharge communication tool. Measuring and comparing the alignment between patient and provider perspectives enables the unveiling of true shared understanding in terms of discharge education [14].

In Hong Kong, the provision of discharge medication information, particularly regarding side effects and warnings, was found to be suboptimal, according to a regular patient experience survey [15]. In 2017, the Hospital Authority developed a computer-generated written medication reminder called the postdischarge information summary (PDIS) to address this issue [16]. The key components of the PDIS were co-designed by a multidisciplinary program team consisting of government officials, clinicians, quality and safety representatives, and technology experts. The first component includes a salient medication reminder, a computer-based drug database encompassing 58 prescribed medications for local older adult patients, and 235 most pertinent side effects and warning items. This database underwent validation through 3 rounds of Delphi expert consensus meetings [17]. The second component comprises a list of follow-up appointments across all Hong Kong public hospitals. The PDIS system generates personalized information by integrating into the electronic health record (EHR). During discharge, physicians or nurses are required to print the written summary through the PDIS system and distribute it to discharged patients or their caregivers, along with a detailed explanation of its contents. No teach-back was required at the moment of program introduction. A comparison of the discharge communication workflow between usual practice and PDIS-incorporated practice is shown in Multimedia Appendix 1.

The objective of this study is to evaluate this EHR-based discharge communication tool by examining and contrasting the perspectives of both older adult patients and health care providers.

Methods

Study Design

This study comprises a self-administered cross-sectional staff survey and a pre-post cross-sectional patient survey. The pre-post patient surveys were conducted among 2 different patient groups. The STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guideline [18] was used to strengthen the reporting process (Checklist 1).

Setting and Sampling

The study involved 4 piloting public tertiary hospitals representing 3 out of 7 geographical clusters of Hong Kong. The PDIS was introduced in a phased manner within the geriatric and medicine department in January 2018 [16]. For the staff survey, all physicians and nurses involved in the PDIS implementation were invited to participate in the study. The surveys were conducted at least 6 months after the PDIS implementation, which spanned from August 2018 to June 2019. Paper-based promotion leaflets, invitation letters, and questionnaires were distributed through designated coordinators in each hospital. This survey was conducted anonymously and on a voluntary basis.

For the patient survey, the sample size was determined based on the inpatient discharge statistics for patients aged 65 years or older in 2015, as provided by the Hospital Authority. In order to achieve a precision level of ±4% with a 95% CI, a minimum of 1450 respondents was required for pre-post rounds. Assuming a 50% response rate, at least 2900 patients were randomly selected from the discharge records for each round. Responses from caregivers acting as surrogates were accepted if patients were unable to respond independently. Readmitted cases and day patients were excluded. Within 14 days of their discharge, patients were contacted by telephone. The pre-post survey was conducted from June to December 2017 and May to December 2018, accordingly.

The staff survey used English and the patient version used Chinese. We used English in staff survey because English is their working language and they are proficient in English. Furthermore, staff surveys are typically conducted in English in Hong Kong, so it was assumed that participants would feel comfortable with this language. To ensure language did not pose a barrier to participation or affect the responses collected, we decided to use Chinese for the patient survey. We do not anticipate any language concordance issues because the researchers who were responsible for collecting patient responses were well-trained before the survey to ensure they can correctly convey the meaning of the questions.

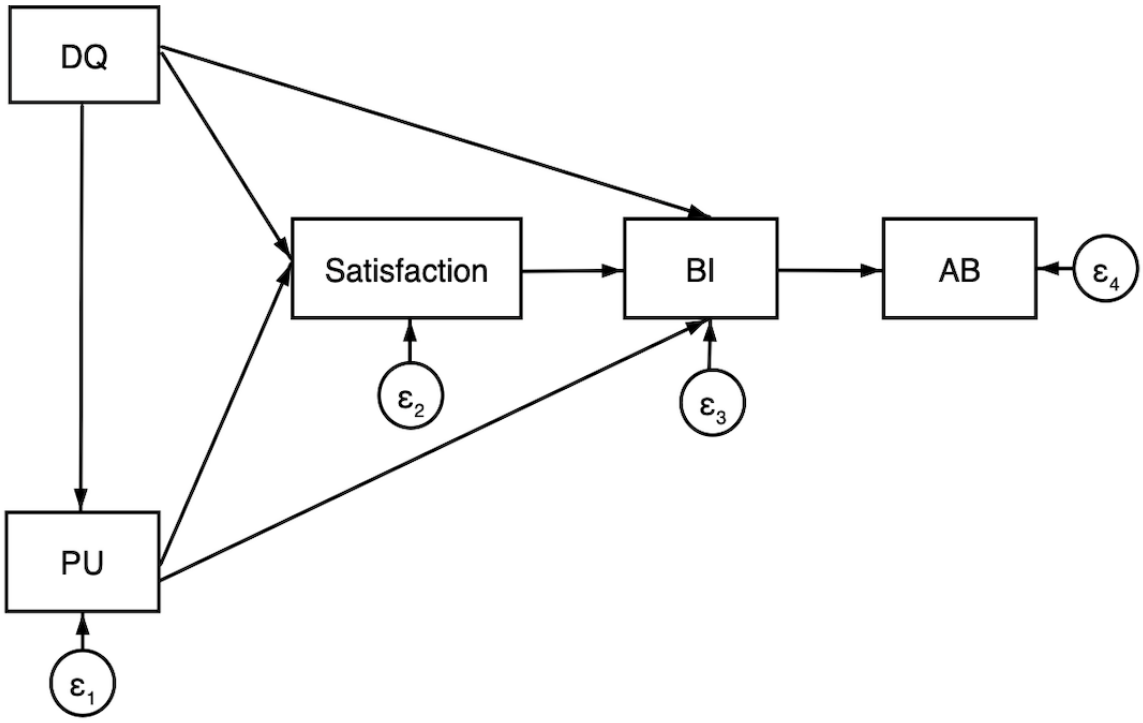
Theoretical Framework

The staff survey collected information on providers' practicing behaviors and user experience, adapted by the technology acceptance model (TAM) [19]. TAM has been designed to investigate why individuals adopt a specific technology and has been widely used in different settings [20].

According to the TAM, the perceived usefulness of the technology can impact users' behavior intention, which can be a determinant of users' actual behaviors. We added another domain named design quality to capture providers' perceived information quality of the tool. We also measured the overall satisfaction of the PDIS experience. Existing literature suggests that design quality can significantly influence staff adoption and implementation of the technologies [21]. The traditional TAM component of "Perceived Ease of Use" was not fully applicable in this context, as our primary aim was to compare the perceptions of patients and staff regarding their user experiences and perceptions of the PDIS. However, "Perceived Ease of Use" tends to be more relevant to staff, given that staff interact with the information system while patients interact with a paper-based version rather than a technology. Consequently, we adopted "Design Quality" as a domain to facilitate a meaningful comparison of perspectives between these 2 groups. In addition, we sought to streamline our survey to enhance response rates, particularly among staff who may have limited time to participate. Thus, we decided to use design quality as a domain in place of perceived ease of use.

We hypothesized that (1) design quality would have an impact on perceived usefulness, satisfaction, and behavior intention; (2) perceived usefulness would impact satisfaction and behavior intention; (3) satisfaction would impact behavior intention; and (4) behavior intention would impact actual behavior (Figure 1).

Figure 1. Conceptual framework of the factors impacting implementation fidelity. AB: actual behavior; BI: behavior intention; DQ: design quality; PU: perceived usefulness; "ε1", "ε2", "ε3", and "ε4" were the residual errors.



Measurements

The Staff Survey

Perceived Usefulness

In total, 4 items were used to measure perceived usefulness: (1) "PDIS supports my medication education," (2) "PDIS enhances my communication with patients or caregivers," (3) "PDIS enhances my job efficiency," and (4) "I find PDIS useful in my job." The Cronbach α for this domain was 0.97, indicating the good reliability. A Likert scale ranging from 0 ("strongly disagree") to 10 ("strongly agree") was used for these 4 items.

Design Quality

Design quality was measured using 3 items developed by the research team: (1) "The information provided by the PDIS is clear," (2) "The information provided by the PDIS is sufficient," and (3) "The information provided by the PDIS is useful to your patients or caregivers." The Cronbach α for this domain was 0.92, indicating good reliability. A 10-point Likert scale was used to measure these three items ranging from 0 ("strongly disagree") to 10 ("strongly agree").

Behavior Intention

One item, which was investigator team constructed, was adopted to measure behavior intention: "I am willing to use PDIS to deliver patient discharge information in the future." The 10-point

Likert scale was used (0: strongly disagree to 10: strongly agree) for this item.

Satisfaction

Satisfaction was measured using a single item: "How would you rate your experience of delivering patient discharge information using PDIS?" A 10-point Likert scale ranging from 0 ("very poor") to 10 ("very good") was used for this question.

Actual Behaviors

The actual behavior was defined as the implementation fidelity and measured by the frequency of distributing, reading, and explaining the PDIS to patients and caregivers with a 3-level ordinal scale (never, sometimes, and always). A composite score for actual behavior was created by converting the ordinal scale for each behavior into to a 0, 1, and 2 scale and summing it up across the 3 practicing behaviors (distribution, read, and explanation) to generate a total score.

Qualitative Comments

Free-text fields were provided to solicit provides' comments on their PDIS experiences. Respondents were prompted with the following question: "Is there anything you would like to tell us about the PDIS (eg, things that were particularly good, areas for improvement, or any other comments)?"

Demographic Information

Demographic information was collected in the final section of the survey. For example, we asked their self-reported gender, age, working experience in years, and professional role (eg, interns, residents, specialists, associate consultants, and consultants for physicians; enrolled nurses, registered nurses, advanced practice nurses, and ward managers for nurses).

The Patient Survey

The pre-post patient surveys consist of the following sections.

The Patient Experience

The items were drawn from the validated local assessment tool, the Short-form Hong Kong Inpatient Experience Questionnaire (SF-HKIEQ) [22], soliciting patient agreement on the clarity, adequacy, and usefulness of discharge medication information, including side effects and warnings, using an ordinal scale (yes, to some extent, or no) or Likert scale of 0 - 10 from strongly disagree to strongly agree; the Cronbach α is 0.87, indicating good reliability. The overall satisfaction of the PDIS experience was also solicited. A 0 - 10 Likert scale of strongly disagree/very bad to strongly agree/very good was applied to measure the items. In the patient survey, consistent with the previous analysis approach for the SF-HKIEQ, questions with an ordinal scale were converted to a 0, 5, and 10 scale and aggregated to calculate the mean and SD.

Self-Reported Medication-Taking Behavior

This section was developed based on the 4-item Morisky, Green, Levine (MGL) scale [23] and relevant studies [24]. Furthermore, 1 item was adopted from the MGL scale: "Ever forget to take medicines?" In addition, 2 items were investigator-constructed: "Whether you were compliant with their medication regimen" and "Whether you have ever experienced medication side effects." These 3 items were measured by binary responses (yes or no).

Design Quality

The patient postsurvey includes the same question of the design quality domain of the staff survey, allowing comparative analysis: (1) "The PDIS information is clear to me or patients," (2) "The PDIS information is enough to me or patients," and (3) "The PDIS information is useful to self-care or care for patients." A 10-point Likert scale was used (0: "strongly disagree" to 10: "strongly agree"). The Cronbach α was 0.94, indicating a high reliability.

Free-Text Field

A free-text section was added to solicit patients' comments on the PDIS experiences. Similarly with staff survey, we asked patients "what information do you need that was not provided by the PDIS, and any other comments?"

Patient Characteristics

Patient characteristics, such as their age, self-reported gender, education level (primary, secondary, college, and above), whether received the government subsidy, the comprehensive social security assistance designed for people whose income is not sufficient to meet basic needs), living arrangements (living alone or living with others), chronic conditions (including heart disease, hypertension, type 2 diabetes, and cancer), and self-reported quality of life using visual analog scale of the EQ-5D-5L Hong Kong [25], were also asked.

Data Analysis

The statistical analysis was performed utilizing R version 4.0.5 (R Project for Statistical Computing) and Stata version 18 for Mac (StataCorp). Provider and patient demographic information were summarized as means and percentages using descriptive statistical analysis. The staff survey and patient survey were analyzed separately. For the staff survey, subgroup analysis was conducted to evaluate the differences among physicians and nurses. The practicing behavior frequency and PDIS experiences were compared using the Pearson chi-square test and Mann-Whitney *U* test. The Mann-Whitney *U* test was used due to the outcome variables were skewed. In order to understand the relationship between the determinants of the behaviors like design quality, perceived usefulness, behavior intention, satisfaction of the PDIS, and the actual behavior, as well as the relationship between different determinants, the covariance-based structural equation modeling (CB-SEM) was applied. The goodness of fit index, comparative fit index (CFI), root mean square error approximation (RMSEA), Tucker-Lewis index (TLI), and standard root mean squared residual (SRMSR) were used to evaluate the goodness-of-fit of the model. The direct effects, indirect effects, and total effects between the core variables were assessed using a bootstrapping approach (n=2000). Standardized path coefficients (β) were used to estimate the path relationships. Statistical significance was set at P=.05.

For the patient survey, changes in patient experience regarding medication information before and after PDIS implementation were compared using the Mann-Whitney *U* test due to the

outcome variables were skewed. The difference between self-reported side effects encounters and compliance was assessed using the Pearson chi-square test.

To identify the disparity between staff and patient. the shared questions related to PDIS experience (eg, information clarity, adequacy, and usefulness, and the overall experience of the PDIS) in the staff and patient postsurveys were compared using the Mann-Whitney *U* test. For the qualitative comments from both staff and patient, thematic synthesis [26] was applied to identify and compare the common themes in free-text comments for the PDIS program between patients and providers. Furthermore, 2 coders independently coded all the utterances of patients and staffs to generate the initial codes to achieve consensus on coding. The disagreement were formally resolved through several discussions with a senior researcher experienced in qualitative study. After that, the inductive analysis was performed by one coder to generalize initial codes into overarching themes. The initial summary themes were developed and discussed with the senior research to reach consensus. The final list of themes was reviewed and consented by the research team. The theme list were identical for patient and staff and the frequency of each theme were counted in order to compare the pattern for patients and staffs.

Ethical Considerations

Ethical approval has been provided by the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee in compliance with the Declaration of Helsinki (CREC 2019.436). For the staff survey, implied consent was applied as participating staff members returned the completed questionnaires to the research team. For the patient survey, verbal consent was obtained from patients over the phone before the survey. All data were deidentified to protect participants' privacy and confidentiality.

Results

Comparative Analysis

A total of 1375 providers completed the survey with a 76% response rate, comprising 966 (72%) female participants, 650 (50%) participants aged 18 - 29 years old, 595 (50%) participants with 0 - 5 years of working experience, and 1216 (88%) nurses (Table 1).

Table . Demographic information of health care providers.

Characteristics		Total (N=1375)	Doctors (n=159)	Nurses (n=1216)
Sex, n (%) [a]				
	Female	966 (72)	55 (35)	911 (77)
	Male	382 (28)	103 (65)	279 (23)
Age (years), n (%) [b]				
	18-29	650 (50)	52 (34)	598 (52)
	30-39	349 (27)	50 (32)	299 (26)
	40-49	206 (16)	32 (21)	174 (15)
	50-59	86 (6.6)	21 (14)	65 (5.7)
	>59	5 (0.4)	0 (0)	5 (0.4)
Working experience (years), n (%) [c]				
	0-5	595 (50)	46 (34)	549 (53)
	6-10	278 (24)	25 (18)	253 (24)
	11-15	94 (8.0)	23 (17)	71 (6.8)
	16-20	125 (11)	18 (13)	107 (10)
	20-25	42 (3.6)	13 (9.6)	29 (2.8)
	26-30	37 (3.1)	10 (7.4)	27 (2.6)
	>30	9 (0.8)	1 (0.7)	8 (0.8)

[a]31 participants excluded from analysis due to missing information.

[b]82 participants excluded from analysis due to missing information.

[c]185 participants excluded from analysis due to missing information.

From the patient side, we received 2353 valid responses, including 1109 (47%) and 1244 (53%) responses collected through the pre- and postsurveys, respectively. The response rate was 55.5% for the presurvey and 59.4% for the postsurvey. The demographic composition was similar between the pre- and postsurvey groups, except that 6.4% (presurvey group: 853/1106; postsurvey group: 1023/1226) more participants were receiving the government subsidy in the postsurvey group (P<.001) (Table 2).

Table . Demographic information of older adult patients.

Characteristics		Total (N=2353)	Presurvey (n=1109)	Postsurvey (n=1244)
Age (years), mean (SD)		77.48 (7.98)	77.54 (8.00)	77.65 (7.93)
Sex, n (%)				
	Female	1070 (45.5)	517 (46.6)	553 (44.5)
	Male	1283 (54.5)	592 (53.4)	691 (55.5)
Education, n (%) [a]				
	≤Primary	1461 (62.8)	695 (63.4)	946 (62.3)
	Secondary	702 (30.2)	341 (31.1)	361 (29.4)
	≥College	162 (7)	60 (5.5)	102 (8.3)
Living status, n (%) [b]				
	Living alone	292 (12.4)	138 (12.4)	154 (12.4)
	Living with others	2056 (87.6)	971 (87.6)	1085 (87.6)
Government subsidy, n (%) [c]				
	Yes	1880 (80.5)	853 (77.1)	1023 (83.5)
	No	456 (19.5)	253 (22.9)	203 (16.5)
Heart diseases, n (%) [d]				
	Yes	844 (36.1)	398 (35.9)	446 (36.3)
	No	1492 (63.9)	710 (64.1)	782 (63.7)
Hypertension, n (%) [e]				
	Yes	1380 (59.1)	671 (60.6)	709 (57.7)
	No	956 (40.9)	437 (39.4)	519 (42.3)
Type 2 diabetes, n (%) [f]				
	Yes	754 (32.3)	358 (32.3)	396 (32.2)
	No	1582 (67.7)	750 (67.7)	832 (67.8)
Cancer, n (%) [g]				
	Yes	156 (6.7)	85 (7.7)	71 (5.8)
	No	2181 (93.3)	1023 (92.3)	1158 (94.2)
Length of stay (day), n (%) [h]				
	0-3	1256 (53.7)	592 (53.6)	664 (53.7)
	4-7	639 (27.3)	296 (26.8)	343 (27.7)
	>7	446 (19.1)	216 (19.6)	230 (18.6)
EQ-5D-VAS, mean (SD)[i]		66.61 (18.51)	68.37 (17.39)	65.04 (19.33)
Discharge day, n (%)				
	Weekday	1916 (81.5)	909 (82)	1007 (80.9)
	Weekend	437 (18.5)	200 (18)	237 (19.1)

[a]29 participants excluded from analysis due to missing information.
[b]5 participants excluded from analysis due to missing information.
[c]17 participants excluded from analysis due to missing information.
[d]17 participants excluded from analysis due to missing information.
[e]17 participants excluded from analysis due to missing information.
[f]17 participants excluded from analysis due to missing information.
[g]16 participants excluded from analysis due to missing information.

[h]12 participants excluded from analysis due to missing information.

[i]109 participants excluded from analysis due to missing information.

The comparative evaluation showed that patients consistently provided significantly higher ratings for their overall PDIS satisfaction compared with providers (mean 8.28, SD 1.60 vs mean 6.29, SD 1.88, respectively; *P*<.001) (Figure 2). Specifically, patients reported higher ratings for information clarity (mean 8.58, SD 2.50 vs mean 6.54, SD 1.86, respectively; *P*<.001) and usefulness of the PDIS (mean 8.14, SD 2.46 vs mean 6.50, SD 2.02, respectively; *P*<.001). On the contrary, patients were inclined to receive more information through the PDIS (mean 5.33, SD 1.35 vs mean 6.14, SD 2.08, respectively; *P*<.001).

Figure 2. Comparison of patient and health care providers' postdischarge information summary experiences. *P* value was obtained from the Mann-Whitney *U* test. **P*<.001.

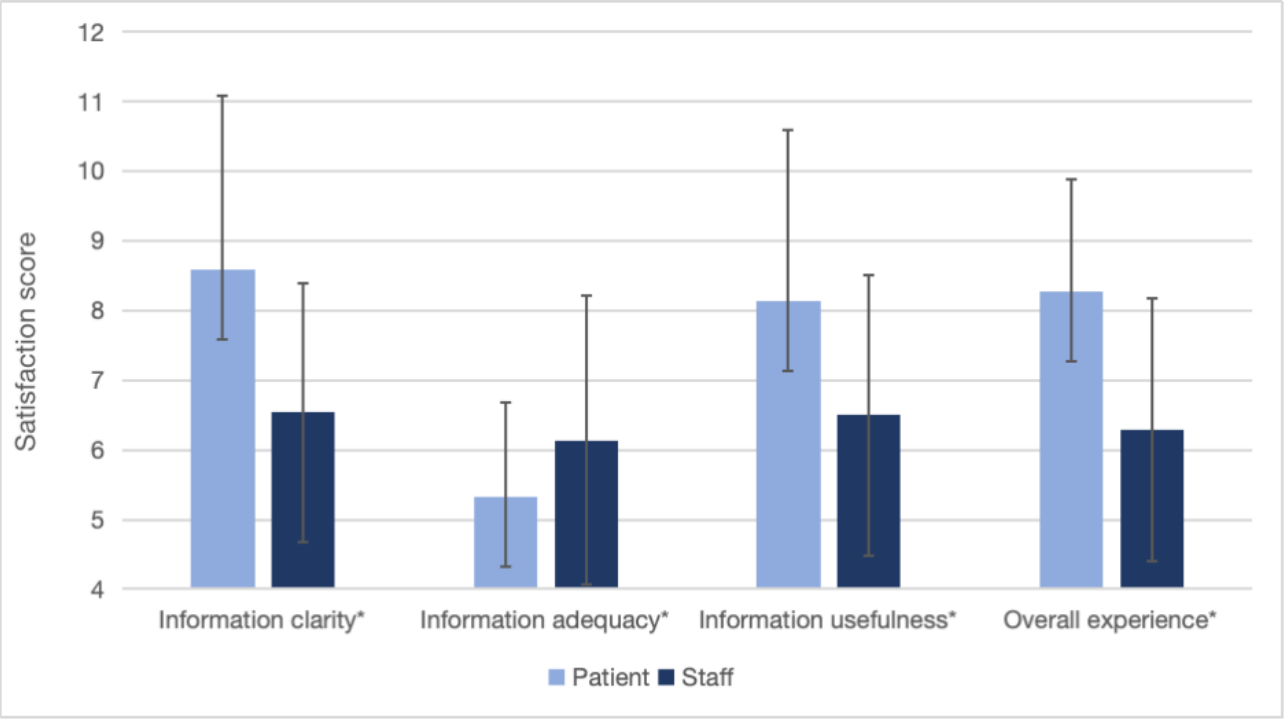
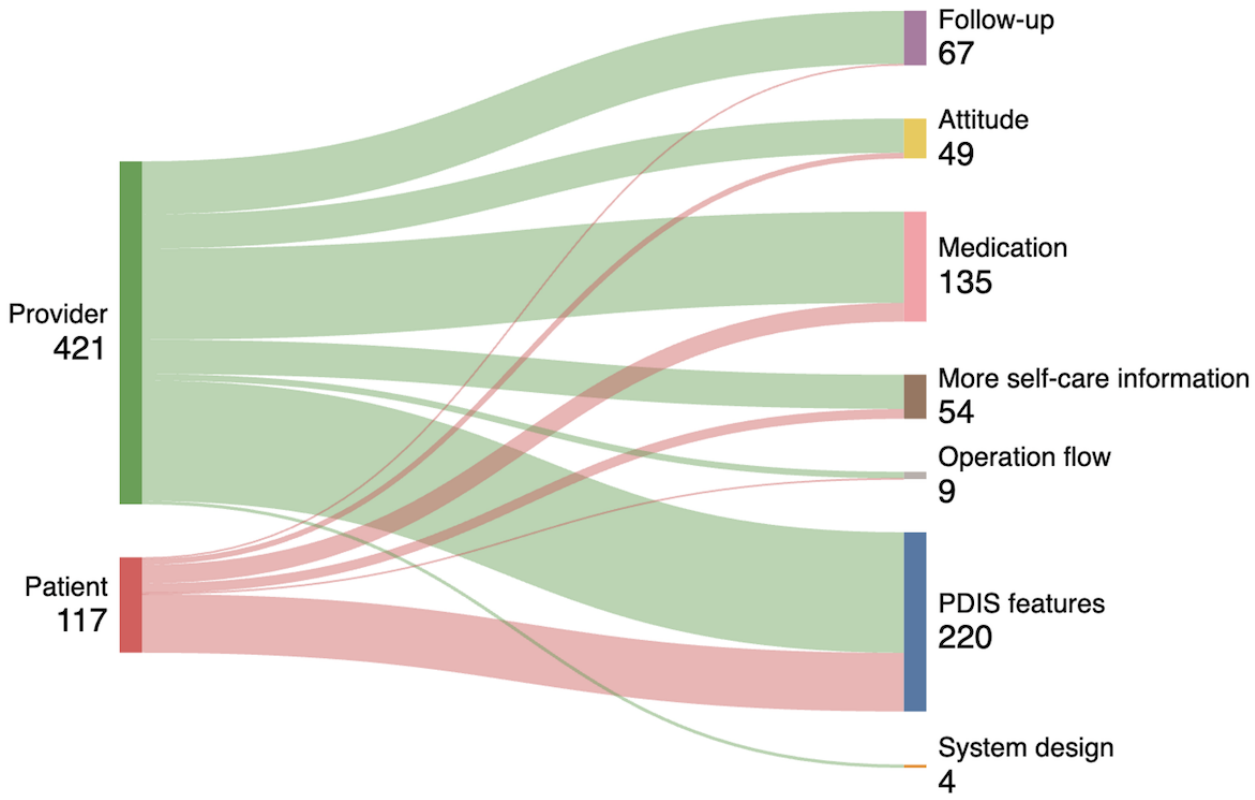


Figure 3 displays the similarities and differences between patients and providers regarding their comments on PDIS experiences. A total of 538 comments were received (421 were from providers and 117 were from patients). The most frequently commented aspect was PDIS features for both providers (147/421, 35%) and patients (73/117, 62%). However, the specific area of the feature was different. Providers emphasized the need for broader coverage of the drug databases (60/147, 41%) and the lack of multiple language versions (59/147, 40%), while patients' concerns revolve around the inconvenience of medication names and follow-up information being in English (45/73, 61%), the discomfort with medical jargon (9/73, 13%), and font size (7/73, 10%). In addition, providers frequently commented on the content of the PDIS form. For example, 27% (114/421) of the comments were related to the medication listed on the PDIS and emphasized the need for additional medication details such as medication changes, indications, and instructions.

Figure 3. Comparison of the qualitative comments on the postdischarge information summary between patients and health care providers. PDIS: postdischarge information summary.



Staff Survey

Analysis of the practicing behaviors (Table 3) revealed that 72.1% (n=922) of the providers reported being able to consistently distribute the PDIS form to patients or caregivers, whereas 56% (n=667) stated they could consistently explain its content. Subgroup analysis demonstrated significant variations across different roles. Regarding distribution, 78% (n=915) of nurses reported always doing so, compared with 6.4% (n=7) of doctors (P<.001). Similarly, 57% (n=666) of nurses reported always explaining the content, compared with 4% (n=1) of doctors (P<.001).

Table . Health care providers' practicing behaviors and experience of the postdischarge information summary.

Items		Total (N=1375)	Doctors (n=159)	Nurses (n=1216)	*P* value[a]
PDIS[b] implementation behavior, n (%)					
Distribute		1234 (89.7)	109 (68.6)	1170 (96.2)	<.001
	Always	922 (72)	7 (6.4)	915 (78)	
	Sometimes	261 (20)	17 (16)	244 (21)	
	Never	96 (7.5)	85 (78)	11 (0.9)	
Read		1190 (86.5)	26 (16.4)	1154 (94.9)	<.001
	Always	730 (61)	5 (19)	725 (62)	
	Sometimes	425 (36)	16 (62)	409 (35)	
	Never	35 (2.9)	5 (19)	30 (2.6)	
Explain		1190 (86.5)	26 (16.4)	1164 (95.7)	<.001
	Always	667 (56)	1 (3.8)	666 (57)	
	Sometimes	475 (40)	11 (42)	464 (40)	
	Never	48 (4)	14 (54)	34 (2.9)	
Perceptions of PDIS experiences, mean (SD)					
Perceived design quality					
	PDIS information is clear to your patients or careers	6.54 (1.86)	6.50 (1.61)	6.54 (1.86)	.67
	PDIS information is adequate to your patients or careers	6.14 (2.07)	6.24 (1.61)	6.14 (2.08)	.93
	PDIS information is useful for patients or careers	6.50 (2.02)	6.69 (1.81)	6.50 (2.02)	.70
Perceived usefulness					
	PDIS supports my medication education to patients or careers	6.35 (2.07)	6.08 (1.81)	6.36 (2.07)	.32
	Patient-provider communication becomes more effective with PDIS	6.21 (2.05)	6.00 (1.55)	6.21 (2.06)	.34
	PDIS enhances my job efficiency	5.97 (2.20)	6.08 (1.72)	5.97 (2.21)	.83
	PDIS is useful in my job	6.03 (2.16)	6.12 (1.68)	6.02 (2.17)	.97
Behavior intention					
	I am willing to use PDIS	6.09 (2.18)	6.12 (1.63)	6.09 (2.19)	.72
Overall satisfaction					
	Overall rating of PDIS user experiences	6.29 (1.88)	6.56 (1.51)	6.28 (1.88)	.75

[a]*P* value was obtained from the chi-square test and Mann-Whitney *U* test.

[b]PDIS: postdischarge information summary.

Physicians and nurses indicated moderate satisfaction with the design quality and perceived usefulness of the PDIS to their work, as reflected by mean agreement scores ranging between 5.96 and 6.54 (Table 3). The subgroup analysis did not identify

any differences between professional roles regarding user experiences. The CB-SEM analysis (Figure 4) showed that design quality significantly impacted their perceived usefulness (β coefficient=0.96, 95% CI 0.90-1.01) and behavior intention (β coefficient=0.14, 95% CI 0.06-0.21). In addition, perceived usefulness had a significant impact on behavior intention (β coefficient=0.48, 95% CI 0.26-0.70). Furthermore, behavior intention had a significant impact on the actual behavior (β coefficient=0.16, 95% CI 0.10-0.23). In addition, satisfaction can be significantly impacted by the design quality (β coefficient=0.24, 95% CI 0.08-0.40) and perceived usefulness (β coefficient=0.57, 95% CI 0.43-0.71). The structural equation modeling (SEM) model presents a good fit overall, with all indicators exceeding the recommended thresholds (Table 4). The results of the indirect effects and total effects can be found in Multimedia Appendices 2 and 3. In total, 3 mediating pathways were identified: (1) an indirect pathway from design quality through satisfaction to behavior intention (β coefficient=0.770, 95% CI 0.700 - 0.841, proportion mediated=15.1%); (2) an indirect pathway from perceived usefulness through satisfaction to behavior intention (β coefficient=0.228, 95% CI 0.185 - 0.272, proportion mediated=32.5%); and (3) an indirect pathway from design quality through perceived usefulness to satisfaction (β coefficient=0.544, 95% CI 0.482 - 0.605, proportion mediated=30.8%).

Figure 4. Structural equation modeling for factors impacting the providers' implementation fidelity. AB: actual behavior; BI: behavior intention; DQ: design quality; PU: perceived usefulness; "ε1", "ε2", "ε3", and "ε4" were the residual errors.

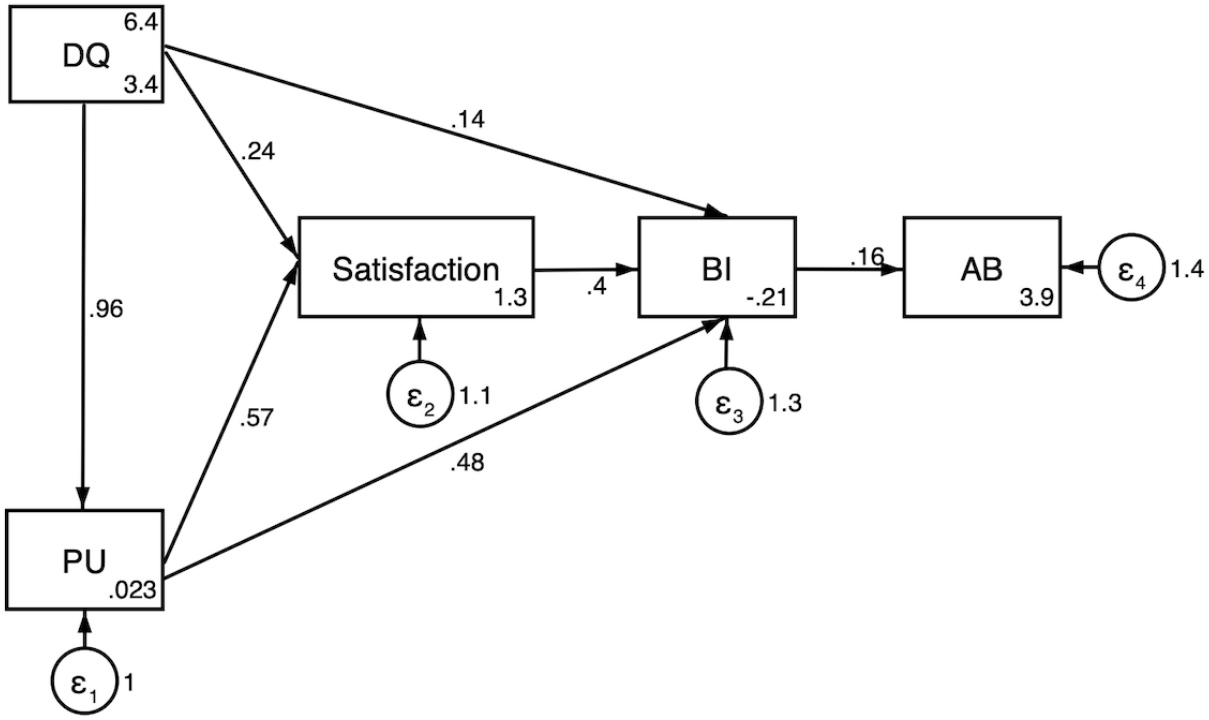


Table . Direct effects for the model[a].

Relationship	Standardized estimates (95% CI)	Remarks
Design quality → Perceived usefulness	0.955 (0.923-0.987)	Supported
Design quality → Satisfaction	0.242 (0.151 - 0.333)	Supported
Perceived usefulness → Satisfaction	0.569 (0.487-0.651)	Supported
Design quality → Behavior intention	0.137 (0.031-0.242)	Supported
Perceived usefulness → Behavior intention	0.476 (0.371-0.582)	Supported
Satisfaction → Behavior intention	0.402 (0.288-0.515)	Supported
Behavior intention → Actual behavior	0.164 (0.131-0.197)	Supported

[a]Model fit: χ_{441}^2=3088.344, *P*<.001; root mean square error approximation=0.052; comparative fit index=0.997; Tucker-Lewis index=0.990; standard root mean squared residual=0.021.

Patient Survey

Table 5 showed significant improvements in patient experience related to overall discharge information clarity (mean 8.18, SD 1.69 vs mean 7.93, SD 1.84, respectively; *P*=.002), adequacy (mean 8.15, SD 1.76 vs mean 7.92, SD 1.93, respectively; *P*=.01), and usefulness (mean 8.26, SD 1.70 vs mean 8.06, SD 1.80, respectively; *P*=.02). In addition, a significant increase was found in information adequacy for both side effects (mean 9.6, SD 2.0 vs mean 8.6, SD 3.4, respectively; *P*<.001) and warnings (mean 9.7, SD 1.8 vs mean 9.2, SD 2.7, respectively; *P*=.004). Only warning information reached statistically significant improvement in clarity (mean 8.77, SD 2.32 vs mean 8.45, SD 2.45, respectively; *P*=.03) and usefulness (mean 8.7, SD 2.45 vs mean 8.44, SD 2.39, respectively; *P*=.03).

Table . Older adult patients' or caregivers' perceptions of discharge medication information and medication-taking behavior between pre- and postsurvey groups.

Items		Total (N=2353)	Presurvey (n=1109)	Postsurvey (n=1244)	*P* value[a]
Perspectives of the discharge medication information					
Clarity, mean (SD)					
	Side effects	8.47 (2.50)	8.31 (2.63)	8.60 (2.39)	.10
	Warning signs	8.61 (2.39)	8.45 (2.45)	8.77 (2.32)	.03
	Overall	8.06 (1.77)	7.93 (1.84)	8.18 (1.69)	.003
Adequacy, mean (SD)					
	Side effects	9.2 (2.8)	8.6 (3.4)	9.6 (2.0)	<.001
	Warning signs	9.4 (2.3)	9.2 (2.7)	9.7 (1.8)	.004
	Overall	8.04 (1.85)	7.92 (1.93)	8.15 (1.76)	.02
Usefulness, mean (SD)					
	Side effects	8.47 (2.46)	8.34 (2.51)	8.57 (2.42)	.12
	Warning signs	8.57 (2.42)	8.44 (2.39)	8.70 (2.45)	.03
	Overall	8.17 (1.75)	8.06 (1.80)	8.26 (1.70)	.01
Medication-taking behaviors					
Self-reported side effects encounter, n (%)		2318 (86.1)	1083 (85.6)	1235 (86.6)	.04
	Yes	237 (10.2)	126 (11.6)	111 (9)	
	No	2081 (89.8)	957 (88.4)	1124 (91)	
Self-reported medication compliance, n (%)		2325 (86.4)	1085 (85.8)	1240 (87)	.39
	Yes	2226 (95.7)	1043 (96.1)	1183 (95.4)	
	No	99 (4.3)	42 (3.9)	57 (4.6)	

[a]*P* value was obtained from the Mann-Whitney *U* test and chi-square test.

No statistically significant difference was found between pre- and postsurvey groups in the percentage of self-reported medication compliance. Notably, the postsurvey group had a significantly lower percentage of self-reported side effects encounters (126/1083, 11.6% vs 111/1235, 9%, respectively; *P*=.04). Among participants who reported encountering side effects, the majority (96/111, 86.4%) still followed the medication instructions as prescribed.

Discussion

Principal Results

This is the first study to compare the perceptions of older adult patients and health care providers regarding the use of large-scale EHR-based discharge communication tools with concordance measures. There was a noticeable difference in ratings between patients and providers, with patients giving higher ratings in terms of design quality and overall experience of this tool. Qualitative comments indicated that patients and providers have different areas of concern regarding this communication tool. Furthermore, from the health care providers' side, inconsistent practicing behaviors were found, which were significantly influenced by the implementation intentions (represented as behavioral intentions in the SEM model), overall satisfaction, design quality, and perceived usefulness of the program. However, from the patients' side, older adults who received the written summary reported improved experiences with discharge information, including information clarity, adequacy, and usefulness.

Comparison With Previous Work

Providers assigned significantly lower scores to the perceived information clarity of the PDIS to their patients than patients themselves. This difference may be due to the providers' concern about the challenges associated with older adults' health literacy [27] and the potential negative consequences of sharing information on side effects, such as patients' anxiety and non-compliance [28-30]. However, the improved patient-reported ratings of information clarity on medication warnings and overall medication information, significantly fewer side effects encounters, and no evidence of patient noncompliance found by our study and others [31] suggested that taking action is no worse than inaction but yields better outcomes, contrary to the biased perception held by staff members [32]. In order to address staff concerns, facilitate their implementation, and fulfill patients' needs, rephrasing risk information by using lay language, shorter sentences, supplementing verbal descriptions with visual aids, and presenting medication benefits along with side effects can be considered [33-35].

Providers rated information usefulness for patients or caregivers lower than patients in this study, further impacting their implementation. The lower beliefs on the value of the communication tool for patient care from the provider side may be due to the beliefs that patients may not effectively follow the advice due to a lack of skills or inability to recall, despite clinicians appropriately delivering the instructions [36]. Therefore, it is suggested that using cognitive aid strategies such as teach-back techniques, repetition, demonstration, and reducing the complexity of the information to enhance patients' capacity to perform self-care tasks and recall of information [10]. The discrepancy can also be attributed to providers' lack of awareness regarding patient needs, which was also reported by previous studies [37]. Our study results, as well as other research [38], suggest that patients view information about medication side effects and warnings as crucial when making decisions about seeking professional assistance. Patient-provider information gap may not only lead to patient dissatisfaction but also levy stress on providers [37]. Therefore, it is important to leverage patient voices as credible sources and build long-term patient-provider relationships to address this gap.

It is important to note that older adults tend to rely more on health care providers and perceive them as trustworthy sources of information, as reflected in higher satisfaction with received medication information among older adults than younger individuals in a previous study [39]. Therefore, it is crucial to address provider-reported barriers in PDIS implementation to ensure a higher level of program satisfaction and implementation fidelity. Our study found that the design quality and perceived usefulness to providers' work can hinder their implementation, which is in accordance with previous studies [40,41]. This can be attributed to the low compatibility of the service with providers' existing workflow [21]. Our study found that providers expressed a preference for enriching the PDIS content with additional medication elements, such as medication changes, indications, and instructions, indicating their priorities for discharge education are not fully met. This suggested that involving frontline implementers as program designers and developers is crucial for program fit, staff self-efficacy, capacity, and commitment [21]. Other than the perceived usefulness and design quality of the program reported in our study, a comprehensive understanding of the complex elements involved in implementation, including context, stakeholders, and organizational factors, is needed [42]. This knowledge facilitates the creation of customized strategies and policies that have a higher likelihood of achieving success [43]. Therefore, conducting implementation research is essential in identifying the barriers and facilitators linked to the implementation of the PDIS to ensure providers' user experiences are optimized, leading to improved patient access to high-quality care and maintaining a high level of patient experiences among the older adult population.

Limitations

This study has several limitations. First, self-reported data may be subject to inaccuracies due to social desirability bias. However, including participants from diverse backgrounds may mitigate this limitation to some extent. Second, patient outcomes were not measured in this study, as our focus was on exploring and comparing patient and provider experiences with technology-based communication tools. Given the positive experiences reported by patients, future research could investigate clinical outcomes to further enhance the evidence base. Third, the pre-post survey design for a patient survey without a control group limits our ability to determine whether the observed changes are directly attributable to the PDIS. The cross-sectional design for the staff survey can limit our ability to establish causal inferences regarding factors influencing providers' inconsistent implementation. Subsequent studies using longitudinal or experimental designs are warranted to understand the causal mechanisms and develop effective strategies to enhance staff performance. Finally, as we did not impose a strict designation of specific professional roles (doctor or nurse) for the tasks involved in PDIS implementation, there may be an underestimation of the effects of the determinants on behavioral intention and actual behaviors.

Conclusions

EHR-based discharge communication tools have the potential to improve the patient experience with discharge information. However, there is a notable difference in user perceptions between patients and providers. This difference may hinder the full benefits of the program for patients. These findings have implications for future research, particularly in implementation research, where barriers and strategies to enhance staff performance can be investigated. In addition, the study provides valuable insights for organizations seeking to improve patient-provider shared understanding of postacute care plans among older adult patients during hospitalization, particularly through technology-based interventions.

Acknowledgments

We are grateful to the steering committee members for their professional advice regarding the implementation of the postdischarge information summary. We acknowledge the assistance of Ms Tang PF. We would also like to thank the HA Head Office and the cluster coordinators for their kind assistance in the logistics for conducting the study.

Authors' Contributions

ELYW, EKY, and KST were involved in the conception and design of the study. DYXW is the lead author and prepared the article. AWLC was involved in data collection. DYXW, ELYW, and AWLC were involved in data analysis. All the authors were involved in the critical revision and the final approval of the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Pre- and postdischarge information summary discharge communication workflow.
[DOCX File, 18 KB - aging_v8i1e60506_app1.docx]

Multimedia Appendix 2
Indirect effect for the structural equation modeling model.
[DOCX File, 15 KB - aging_v8i1e60506_app2.docx]

Multimedia Appendix 3
Total effect of the structural equation modeling model.
[DOCX File, 15 KB - aging_v8i1e60506_app3.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement checklist.
[PDF File, 135 KB - aging_v8i1e60506_app4.pdf]

References

1. Becker C, Zumbrunn S, Beck K, et al. Interventions to improve communication at hospital discharge and rates of readmission: a systematic review and meta-analysis. JAMA Netw Open 2021 Aug 2;4(8):e2119346. [doi: 10.1001/jamanetworkopen.2021.19346] [Medline: 34448868]
2. Davis TC, Wolf MS, Bass PF 3rd, et al. Literacy and misunderstanding prescription drug labels. Ann Intern Med 2006 Dec 19;145(12):887-894. [doi: 10.7326/0003-4819-145-12-200612190-00144] [Medline: 17135578]
3. Sanderson BK, Thompson J, Brown TM, Tucker MJ, Bittner V. Assessing patient recall of discharge instructions for acute myocardial infarction. J Healthc Qual 2009;31(6):25-33. [doi: 10.1111/j.1945-1474.2009.00052.x] [Medline: 19957461]
4. Wolf MS, Davis TC, Shrank W, et al. To err is human: patient misinterpretations of prescription drug label instructions. Patient Educ Couns 2007 Aug;67(3):293-300. [doi: 10.1016/j.pec.2007.03.024] [Medline: 17587533]
5. Sweileh WM, Zyoud SH, Abu Nab'a RJ, et al. Influence of patients' disease knowledge and beliefs about medicines on medication adherence: findings from a cross-sectional survey among patients with type 2 diabetes mellitus in Palestine. BMC Public Health 2014 Jan 30;14:94. [doi: 10.1186/1471-2458-14-94] [Medline: 24479638]
6. Al-Qazaz HK, Sulaiman SA, Hassali MA, et al. Diabetes knowledge, medication adherence and glycemic control among patients with type 2 diabetes. Int J Clin Pharm 2011 Dec;33(6):1028-1035. [doi: 10.1007/s11096-011-9582-2] [Medline: 22083724]
7. Husson O, Mols F, van de Poll-Franse LV. The relation between information provision and health-related quality of life, anxiety and depression among cancer survivors: a systematic review. Ann Oncol 2011 Apr;22(4):761-772. [doi: 10.1093/annonc/mdq413] [Medline: 20870912]
8. Koehler BE, Richter KM, Youngblood L, et al. Reduction of 30-day postdischarge hospital readmission or emergency department (ED) visit rates in high-risk elderly medical patients through delivery of a targeted care bundle. J Hosp Med 2009 Apr;4(4):211-218. [doi: 10.1002/jhm.427] [Medline: 19388074]
9. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. N Engl J Med 2009 Apr 2;360(14):1418-1428. [doi: 10.1056/NEJMsa0803563] [Medline: 19339721]
10. Menichetti J, Lie HC, Mellblom AV, et al. Tested communication strategies for providing information to patients in medical consultations: A scoping review and quality assessment of the literature. Patient Educ Couns 2021 Aug;104(8):1891-1903. [doi: 10.1016/j.pec.2021.01.019] [Medline: 33516591]

11. Newnham H, Barker A, Ritchie E, Hitchcock K, Gibbs H, Holton S. Discharge communication practices and healthcare provider and patient preferences, satisfaction and comprehension: a systematic review. Int J Qual Health Care 2017 Oct 1;29(6):752-768. [doi: 10.1093/intqhc/mzx121] [Medline: 29025093]
12. Motamedi SM, Posadas-Calleja J, Straus S, et al. The efficacy of computer-enabled discharge communication interventions: a systematic review. BMJ Qual Saf 2011 May;20(5):403-415. [doi: 10.1136/bmjqs.2009.034587] [Medline: 21262793]
13. Arditi C, Rège-Walther M, Wyatt JC, Durieux P, Burnand B. Computer-generated reminders delivered on paper to healthcare professionals; effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2012 Dec 12;12:CD001175. [doi: 10.1002/14651858.CD001175.pub3] [Medline: 23235578]
14. Street RL. Information giving, managing, and understanding in clinical encounters. Patient Educ Couns 2021 Aug;104(8):1831-1833. [doi: 10.1016/j.pec.2021.05.027] [Medline: 34148629]
15. Wong ELY, Coulter A, Cheung AWL, Yam CHK, Yeoh EK, Griffiths SM. Patient experiences with public hospital care: first benchmark survey in Hong Kong. Hong Kong Med J 2012 Oct;18(5):371-380. [Medline: 23018064]
16. Wong EL, Tang KS, Dong D, et al. Evaluation of the implementation of information system for postdischarge with the theoretical domains framework by healthcare professionals: a multistage design with qualitative inquiry and Delphi expert discussion protocol. BMJ Open 2021 Dec 10;11(12):e046081. [doi: 10.1136/bmjopen-2020-046081] [Medline: 34893482]
17. Wong ELY, Tang KS, Cheung AWL, et al. Development of salient medication reminders to facilitate information transfer during transition from inpatient to primary care: the Delphi process. BMJ Open 2021 May 18;11(5):e041336. [doi: 10.1136/bmjopen-2020-041336] [Medline: 34006537]
18. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
19. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
20. Dehghani H, Mashhadi A. Exploring Iranian English as a foreign language teachers' acceptance of ChatGPT in English language teaching: extending the technology acceptance model. Educ Inf Technol 2024 Oct;29(15):19813-19834. [doi: 10.1007/s10639-024-12660-9]
21. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
22. Wong ELY, Coulter A, Hewitson P, et al. Patient experience and satisfaction with inpatient service: development of short form survey instrument measuring the core aspect of inpatient experience. PLoS One 2015;10(4):e0122299. [doi: 10.1371/journal.pone.0122299] [Medline: 25860775]
23. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [doi: 10.1097/00005650-198601000-00007] [Medline: 3945130]
24. McConachie SM, Berri D, Konja J, Giuliano CA. Evaluation of adverse drug reaction formatting in drug information mobile phone applications. J Med Libr Assoc 2022 Jan 1;110(1):81-86. [doi: 10.5195/jmla.2022.1251] [Medline: 35210966]
25. Wong ELY, Cheung AWL, Wong AYK, Xu RH, Ramos-Goñi JM, Rivero-Arias O. Normative profile of health-related quality of life for Hong Kong general population using preference-based instrument EQ-5D-5L. V H 2019 Aug;22(8):916-924. [doi: 10.1016/j.jval.2019.02.014] [Medline: 31426933]
26. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
27. Shin DW, Roter DL, Cho J, et al. Attitudes toward disclosure of medication side effects: a nationwide survey of Korean patients, caregivers, and oncologists. Psychooncology 2015 Nov;24(11):1398-1406. [doi: 10.1002/pon.3807] [Medline: 25970704]
28. Curtis LM, Mullen RJ, Russell A, et al. An efficacy trial of an electronic health record-based strategy to inform patients on safe medication use: The role of written and spoken communication. Patient Educ Couns 2016 Sep;99(9):1489-1495. [doi: 10.1016/j.pec.2016.07.004] [Medline: 27444235]
29. McGrath JM. Physicians' perspectives on communicating prescription drug information. Qual Health Res 1999 Nov;9(6):731-745. [doi: 10.1177/104973299129122243] [Medline: 10662256]
30. Tarn DM, Wenger A, Good JS, Hoffing M, Scherger JE, Wenger NS. Do physicians communicate the adverse effects of medications that older patients want to hear? Drugs Ther Perspect 2015 Feb;31(2):68-76. [doi: 10.1007/s40267-014-0176-7] [Medline: 25750513]
31. Howland JS, Baker MG, Poe T. Does patient education cause side effects? A controlled trial. J Fam Pract 1990 Jul;31(1):62-64. [Medline: 2193996]
32. Spranca M, Minsk E, Baron J. Omission and commission in judgment and choice. J Exp Soc Psychol 1991 Jan;27(1):76-105. [doi: 10.1016/0022-1031(91)90011-T]
33. Längst G, Seidling HM, Stützle M, et al. Factors associated with medication information in diabetes care: differences in perceptions between patients and health care professionals. Patient Prefer Adherence 2015;9:1431-1441. [doi: 10.2147/PPA.S88357] [Medline: 26508840]

34. Hamrosi KK, Raynor DK, Aslani P. Pharmacist, general practitioner and consumer use of written medicine information in Australia: are they on the same page? Res Social Adm Pharm 2014;10(4):656-668. [doi: 10.1016/j.sapharm.2013.10.002] [Medline: 24239213]
35. McCormack L, Craig Lefebvre R, Bann C, Taylor O, Rausch P. Consumer understanding, preferences, and responses to different versions of drug safety messages in the United States: a randomized controlled trial. Drug Saf 2016 Feb;39(2):171-184. [doi: 10.1007/s40264-015-0358-9] [Medline: 26547718]
36. Peimani M, Nasli-Esfahani E, Sadeghi R. Patients' perceptions of patient-provider communication and diabetes care: a systematic review of quantitative and qualitative studies. Chron Illn 2020 Mar;16(1):3-22. [doi: 10.1177/1742395318782378] [Medline: 29895167]
37. Loewenthal JV, Beltran CP, Atalay A, Schwartz AW, Ramani S. "What's Going to Happen?": Internal medicine resident experiences of uncertainty in the care of older adults. J Gen Intern Med 2024 Mar 14. [doi: 10.1007/s11606-024-08720-y] [Medline: 38485878]
38. Grime J, Blenkinsopp A, Raynor DK, Pollock K, Knapp P. The role and value of written information for patients about individual medicines: a systematic review. Health Expect 2007 Sep;10(3):286-298. [doi: 10.1111/j.1369-7625.2007.00454.x] [Medline: 17678517]
39. van Geffen ECG, Philbert D, van Boheemen C, van Dijk L, Bos MB, Bouvy ML. Patients' satisfaction with information and experiences with counseling on cardiovascular medication received at the pharmacy. Patient Educ Couns 2011 Jun;83(3):303-309. [doi: 10.1016/j.pec.2011.04.004] [Medline: 21550196]
40. Ferdousi R, Arab-Zozani M, Tahamtan I, Rezaei-Hachesu P, Dehghani M. Attitudes of nurses towards clinical information systems: a systematic review and meta-analysis. Int Nurs Rev 2021 Mar;68(1):59-66. [doi: 10.1111/inr.12603] [Medline: 32608032]
41. Jimma BL, Enyew DB. Barriers to the acceptance of electronic medical records from the perspective of physicians and nurses:A scoping review. Inform Med Unlocked 2022;31:100991. [doi: 10.1016/j.imu.2022.100991]
42. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3(1):32. [doi: 10.1186/s40359-015-0089-9] [Medline: 26376626]
43. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53. [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]

Abbreviations

CB-SEM: covariance-based structural equation modeling
CFI: comparative fit index
EHR: electronic health record
MGL: 4-item Morisky, Green, Levine
PDIS: postdischarge information summary
RMSEA: root mean square error approximation
SEM: structural equation modeling
SF-HKIEQ: Short-form Hong Kong Inpatient Experience Questionnaire
SRMSR: standard root mean squared residual
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TAM: technology acceptance model
TLI: Tucker-Lewis index

Edited by X Xiang; submitted 14.05.24; peer-reviewed by T Tang; revised version received 25.11.24; accepted 27.11.24; published 29.01.25.

Please cite as:
Wang DY, Wong ELY, Cheung AWL, Tang KS, Yeoh EK
A Comparison of Patient and Provider Perspectives on an Electronic Health Record–Based Discharge Communication Tool: Survey Study
JMIR Aging 2025;8:e60506
URL: https://aging.jmir.org/2025/1/e60506
doi:10.2196/60506

Association of Subjective Cognitive Concerns With Performance on Mobile App–Based Cognitive Assessment in Cognitively Normal Older Adults: Observational Study

Caroline O Nester[1], PhD; Alyssa N De Vito[1], PhD; Sarah Prieto[1], PhD; Zachary J Kunicki[1], PhD; Jennifer Strenger[1,2], MA; Karra D Harrington[3], PhD; Nelson Roque[4], PhD; Martin J Sliwinski[3], PhD; Laura A Rabin[5], PhD; Louisa I Thompson[1], PhD

[1]Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, 222 Richmond St, Providence, RI, United States

[2]Memory and Aging Program, Butler Hospital, Providence, RI, United States

[3]Center for Healthy Aging, Penn State University, University Park, PA, United States

[4]Department of Psychology, University of Central Florida, Orlando, FL, United States

[5]Department of Psychology, Brooklyn College CUNY, Brooklyn, NY, United States

Corresponding Author:
Caroline O Nester, PhD
Department of Psychiatry and Human Behavior, The Warren Alpert Medical School of Brown University, 222 Richmond St, Providence, RI, United States

Abstract

Background: Subjective cognitive concerns (SCCs) may be among the earliest clinical symptoms of dementia. There is growing interest in applying a mobile app–based cognitive assessment to remotely screen for cognitive status in preclinical dementia, but the relationship between SCC and relevant mobile assessment metrics is uncertain.

Objective: This study aimed to characterize the relationship between SCC and adherence, satisfaction, and performance on mobile app assessments in cognitively unimpaired older adults.

Methods: Participants (N=122; $Mean_{age}$=68.85 [SD 4.93] years; $Mean_{education}$=16.85 [SD 2.39] years; female: n=82, 66.7%; White:n=106, 86.2%) completed 8 assessment days using Mobile Monitoring of Cognitive Change (M2C2), an app-based testing platform, with brief daily sessions within morning, afternoon, and evening time windows (24 total testing sessions). M2C2 includes digital working memory, processing speed, and episodic memory tasks. Participants provided feedback about their satisfaction and motivation related to M2C2 upon study completion. SCC was assessed using the Cognitive Function Instrument. Regression analyses evaluated the association between SCC and adherence, satisfaction, and performance on M2C2, controlling for age, sex, depression, and loneliness. Linear-mixed effects models evaluated whether SCC predicted M2C2 subtest performance over the 8-day testing period, controlling for covariates.

Results: SCC was not associated with app satisfaction or protocol motivation, but it was significantly associated with lower rates of protocol adherence (ß=−.20, *P*=.37, 95% CI −.65 to −.02). Higher SCC endorsement significantly predicted worse overall episodic memory performance (ß=−.20, *P*=.02, 95% CI −.02 to −.01), but not working memory or processing speed. There was a main effect of SCC on working memory performance at day 1 (estimate=−1.05, SE=0.47, *P*=.03) and a significant interaction between SCC and working memory over the 8-day period (estimate=0.05, SE=0.02, *P*=.03), such that SCC was associated with initially worse, then progressively better working memory performance.

Conclusions: SCCs are associated with worse overall memory performance on mobile app assessments, patterns of cognitive inefficiency (variable working memory), and mildly diminished adherence across an 8-day assessment period. Findings suggest that mobile app assessments may be sensitive to subtle cognitive changes, with important implications for early detection and treatment for individuals at risk for dementia.

(JMIR Aging 2025;8:e64033) doi:10.2196/64033

KEYWORDS

subjective cognitive concerns; subjective cognitive decline; digital cognitive assessment; mobile app; app-based; preclinical Alzheimer disease; mild cognitive impairment; MCI; preclinical dementia; mobile monitoring of cognitive change; Cognitive Function Instrument; mHealth; mobile health; applications; cognition; assessment; remote; geriatrics; gerontology; aging; memory; older adult; elderly; digital health; mobile phone

Introduction

Subjective cognitive decline (SCD) refers to self-perceived changes in cognitive functioning in the setting of normative performance on objective cognitive measures [1]. Older adults with SCD are more likely to demonstrate cognitive decline over time compared with their peers who do not endorse significant cognitive concerns [1-6]. This association has led many to posit that SCD may be one of the earliest clinical manifestations of Alzheimer disease and related dementias (ADRD) [1,3]. SCD has been linked with critical ADRD neural substrates and biomarkers, including structural [7-9] and functional [4,10] alterations, white matter dysfunction [11-13], and the presence of amyloid (Aβ) and tau [14-17]. However, it is important to acknowledge that most individuals notice changes in cognition as they age (up to 80% of people over the age of 70 years), and that many individuals with SCD do not ultimately progress to dementia [2]. As such, investigating novel markers within SCD may enhance early risk detection in this preclinical population.

Disentangling the relationship between SCD and performance on traditional neuropsychological measures is inherently complex. By definition, individuals with SCD perform in the normal range on objective cognitive tasks [3]. However, some studies have demonstrated that SCD is not without mild cognitive deficits. For example, there is evidence of mild reductions in processing speed, executive functioning, language, and memory in SCD [18-22], and such minor deficits have been correlated with concurrent Alzheimer disease (AD) biomarkers and future clinical progression [23]. There is presently a multitude of approaches in the field and widespread debate about how to best assess SCD, and measure selection may directly impact the association of SCD with concurrent cognition and risk for future cognitive decline [24-27]. It is also essential to recognize that SCD intrinsically reflects a longitudinal change over time [28]; however, traditional neuropsychological evaluations only capture a snapshot of in-the-moment cognitive status. Furthemore, for individuals with a high cognitive baseline, normative performance on objective testing may actually represent a decline in cognitive performance [1]. Novel and highly sensitive cognitive tools are needed to characterize more subtle deficits experienced by individuals with SCD [28].

There is growing interest in using smartphone-based, digital technology to remotely screen and track cognitive functioning in older adults at risk for dementia or those along the AD continuum. Smartphone usage among older Americans is highly prevalent (62% - 81% of people over the age of 60 use a smartphone) [29], and a number of smartphone-based cognitive assessment apps have been developed [30,31]. Such approaches are not only more accessible, convenient, time-effective, and scalable [31-33], but also may be more sensitive to subtle cognitive changes not captured by traditional paper and pencil cognitive assessments [31,32]. Smartphone-based cognitive assessments may also demonstrate superior ecological validity when completed in the individual's lived context with a familiar piece of technology [30,31,34]. Smartphone technology also facilitates high-frequency, repeated cognitive assessment, which allows for the quantification of subtle markers of decline in older adults, such as intraindividual variability and patterns in cognition over time [30-32,35,36]. However, the self-administered, unsupervised nature of smartphone-based cognitive assessments raises concerns related to task adherence, engagement, and potentially confounding factors that may occur within an uncontrolled environment [32].

Particularly relevant to aging research is the limited understanding of how subjective cognitive concerns (SCCs), the core feature of SCD, may be associated with engagement, satisfaction, and performance on such smartphone-based cognitive assessments. Preliminary evidence suggests that smartphone use may be associated with fewer SCCs in older adulthood and possibly constitute a protective factor against cognitive decline [37]; hence, smartphones may represent a promising tool for novel assessments and potential interventions for at-risk populations with self-perceived cognitive decline. Emerging literature suggests that smartphone-based cognitive testing generally has adequate adherence rates (>70%) [32], but little is understood about factors (such as SCC) which may impact engagement or motivation on unsupervised digital tests [38]. These are especially important relationships to disentangle, given the ubiquity of SCC in aging populations [2], high rates of smartphone usage [29], and growing interest in remote assessment to enhance accessibility to cognitive screening. In addition, the relationships between digital cognitive testing and AD biomarkers [39-41], as well as SCC and AD biomarkers [15,17,42], is supported by emerging research, but no studies to date have directly investigated the relationship between SCC, AD biomarkers, and digital smartphone cognitive tasks.

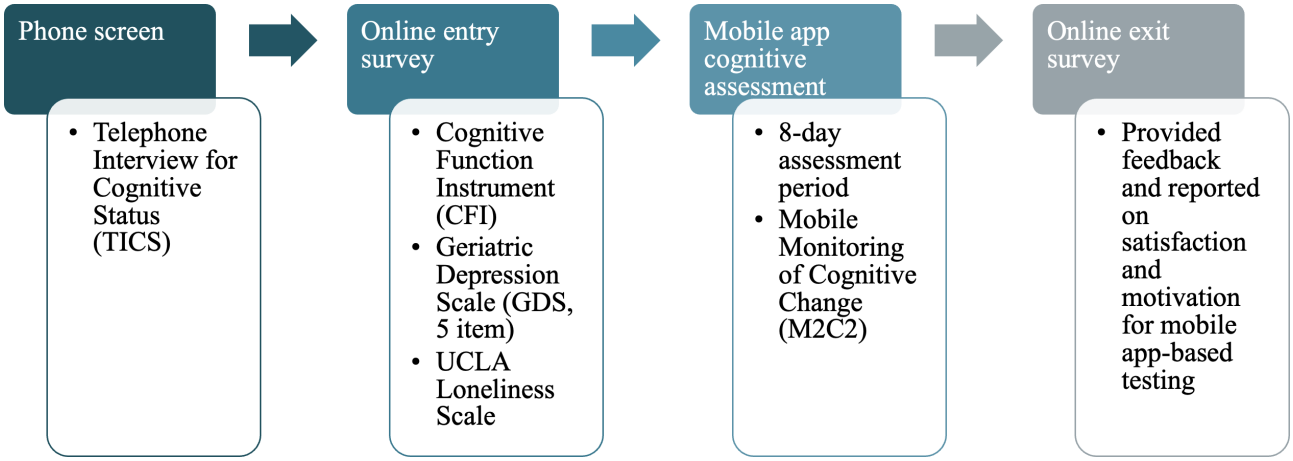
In light of the need to better understand the role of SCC in smartphone-based cognitive assessments in aging populations at risk for AD, the present analysis sought to characterize the relationship between SCC and adherence (aim 1), satisfaction and motivation (aim 2), and performance (aim 3) on a mobile app–based digital cognitive assessment in a sample of cognitively unimpaired older adults. Given the established association between SCC, cerebral Aβ positivity (eg, the hallmark biomarker of AD), and risk for cognitive decline [14-16,23], a secondary aim sought to understand if Aβ positivity moderated the relationship between SCC and performance on the mobile app–based cognitive assessment. Results represent the first step toward quantifying the impact of SCC on smartphone-based cognitive testing in older adult populations, with crucial implications for early detection in those at risk for cognitive decline and dementia.

Methods

Participants, Recruitment, and Procedures

Participants consisted of 122 cognitively unimpaired older adults, between the ages of 60 and 80 years of age, recruited from the Butler Alzheimer's Prevention Registry, a local database of older adults interested in AD research at the Butler Hospital Memory and Aging Program (MAP) [43]. All study procedures were carried out remotely, and detailed recruitment and procedural information has been described previously [38,41,44]. Study procedures are presented in Figure 1.

Figure 1. Study procedures. UCLA: University of California, Los Angeles.



We used a targeted recruitment to enroll individuals with previous amyloid positron emission tomography (PET) data (elevated [Aß+] or nonelevated [Aß-] determined by clinical read by a radiologist) (final n=73) [41] as well as those without PET data. PET results were available if recorded in our Registry as part of previous research participation at the MAP. Individuals (n=256) were invited to participate in the study through an email or phone call. Individuals who consented (n=146) completed an online survey before the initiation of the remote testing protocol, which included various screening measures, including the Cognitive Functioning Index (CFI) [45] (subjective cognitive concern screen), Geriatric Depression Scale 5-item version (GDS-5) [46] (depressive symptoms screen; score range 0 - 5, with higher scores indicating worse depressive symptoms), and the University of California, Los Angeles (UCLA) 3-Item Loneliness Scale [47] (loneliness screen; score range 0 - 6, with higher scores indicating worse loneliness). The modified Telephone Interview for Cognitive Status (TICSm) [48] (objective cognition screen; score range 0 - 50, with lower scores indicating worse cognitive functioning) was also completed before the remote testing protocol by phone, administered by a research assistant. During screening, 23 participants were excluded, and 1 participant was lost to follow up. After the completion of the 8 day remote study protocol, participants (final n=122) were invited to provide feedback (including study satisfaction and motivation) on their experience through an online survey, and a US $20 gift card was provided for compensation.

Inclusion criteria involved unimpaired cognition, daily use of a smartphone, and familiarity with smartphone features. Unimpaired cognition was defined as a TICSm cutoff score of ≥34 [48]. Exclusion criteria included self-reported history of cognitive impairment or dementia, history of neurologic disease or severe mental illness, and physical inability to complete smartphone-based testing.

Subjective Cognitive Concerns

The CFI is a 14-item self-report measure which probes changes in cognitive and functional domains over the past year. The CFI was selected for inclusion in this study as it is a widely used, briefly administered measure of SCC that has been well-validated for used in cognitively healthy and preclinical dementia populations [42,45,49,50]. Items reflect memory concerns, increased use of compensatory strategies, changes in driving, or difficulty managing instrumental activities of daily living. Participants respond with yes (1 point), no (0 points), or maybe (0.5 points), with total scores ranging from 0 to 14 (continuous variable). Higher scores reflect greater SCC. As a sensitivity analysis, we also created a SCC variable that included CFI items that solely assessed cognition (CFI-Cog), by removing the functioning related items to create a refined examination of SCC. CFI-Cog included CFI items 1, 2, 3, 4, 5, 6, 8, 11, and 13. This approach is consistent with a recent study which identified that cognitively focused CFI items, rather than functionally related CFI items, were endorsed at a higher rate and more likely to be associated with Aß+ in cognitively unimpaired older adults [42].

Mobile App–Based Cognitive Assessment

Overview

Remote cognitive tasks were completed using the Mobile Monitoring of Cognitive Change (M2C2) app [51] (more details in Figure 2), a cognitive testing platform developed as part of the National Institute of Aging Mobile Toolbox initiative, with strong previous theoretical, empirical, and psychometric support, including evidence of sensitivity to age and age-related neuropathology [34,38,44]. Android smartphones preloaded with the cognitive assessment app were mailed to participants along with a detailed use guide. Phone functions were locked to prevent the use of features such as web browsing and the camera. Participants completed 3, brief (ie, 3 - 4 minutes) M2C2 sessions each day during morning, afternoon, and evening time windows for 8 consecutive days. Additional sessions could be completed on day 9 as optional or make-up sessions. Staff provided support through phone or email as needed, as described previously [41]. During each M2C2 session, participants completed 3 previously characterized cognitive measures assessing visual working memory (Color Shapes), processing speed (Symbol Match), and episodic memory (Prices). Each task took approximately 60 seconds to complete.

Figure 2. Mobile Monitoring of Cognitive Change (M2C2) app subtests.

Color Shapes

The Color Shapes visual working memory task is a visual array change detection test measuring intraitem feature binding. Participants determine if shapes change color across 2 sequential presentations in which the shape locations change. Performance was measured by discriminability (d-prime) performance calculated from the proportion of correct identifications and proportion of misidentified stimuli [51].

Symbol Match

The Symbol Match processing speed task is a speeded continuous performance test of conjunctive feature search. Participants are asked to identify matching symbol pairs. Performance was measured by the median reaction time to complete the task across all trials (in milliseconds) [51].

Prices

The Prices episodic memory task uses an immediate delayed forced-choice recognition paradigm. Participants incidentally encode 10 grocery item-price pairs for later recall while judging whether or not the item's price is "good." Recall trials begin immediately after the learning trials. Performance was measured by the proportion of correct responses on the 10 recall trials [51].

Study Adherence, Satisfaction, and Motivation

Adherence was quantified as the number of testing sessions completed across the 8-day ambulatory protocol (out of 24 sessions).

Participants' feedback covered various aspects, including their satisfaction with the protocols and their level of motivation to engage with the study tasks. Feedback related to satisfaction was measured with 2 items: "Completing the brain games was fun" and "I became bored with the brain games." Participants rated their responses on a 6-point Likert scale (1=strongly disagree to 6=strongly agree). For the purposes of this study, responses were dichotomized by those who agreed (Likert scale responses 4, 5, and 6) or those who disagreed (Likert scale responses 1, 2, and 3) on these items due to a largely positive response bias (most individuals found the tasks to be fun and not boring).

Feedback regarding motivation was measured on the item: "It was hard to motivate myself to complete the games each time." Participants rated their response on a 5-point Likert scale (1=Agree to 5=Disagree). As above, responses were dichotomized into those who agreed or were neutral (Likert scale responses 1, 2, and 3) and those who disagreed (Likert scale responses 4 and 5) for the purpose of this analysis due to a largely positive response bias (most individuals found themselves to be easily motivated to complete the tasks).

Statistical Analysis

Descriptive statistics were conducted to characterize the sample in terms of demographics (self-reported biological sex, race, ethnicity, and years of education), self-reported symptoms of depression (by the Geriatric Depression Scale 5-item version [46]), loneliness (by the UCLA Loneliness Scale [47]), and SCC (CFI total and CFI-Cog), relevant app-based metrics (ie, adherence, satisfaction and motivation), and objective cognitive performance (ie, visual working memory, processing speed, and episodic memory). Pearson correlation coefficient (for continuous variables) and point biserial correlation (for binary variables) were implemented as appropriate to assess the relationship between SCC and demographic variables, depression, loneliness, and outcomes of interest (eg, adherence, satisfaction, motivation, and performance on app-based cognitive assessment). To address aims 1 - 3 and secondary analyses, SCC (continuous variable) was the main predictor of interest and we adjusted for covariates, including age, sex, depression, and loneliness, in all models. Covariates were selected based on significant association with independent and dependent variables of interest. A priori significance threshold (α) was set at $P \leq .05$. For our moderation analyses, we examined α levels up to $P \leq .10$.

To test aim 1, linear regression models were used to evaluate the association between SCC and adherence (continuous variable) over the 8-day protocol. For the purposes of this analysis, the CFI total score was the main SCC predictor of interest. We present results for CFI-Cog only when they differ from the CFI total score. To test aim 2, logistic regression models were conducted to evaluate the association between SCC and satisfaction (dichotomized self-report items related to level of fun and boredom), and SCC and motivation (dichotomized self-report items related to motivation). Individual regression models were implemented for each motivation and satisfaction outcome of interest. To test aim 3, separate linear regression models were constructed to evaluate the association between SCC and performance on each M2C2 based cognitive assessment (Prices, Color Shapes, and Symbol Search, all continuous variables). Linear-mixed effects models were used to evaluate whether SCC predicted M2C2 subtest performance over time while controlling for age, sex, and loneliness. Finally, as a secondary analysis in a subsample of participants with Aβ PET data, we fit linear regression models with SCC and the interaction of SCC and Aβ PET positivity as predictors, to test the moderation effect of Aβ PET positivity on the association between SCC and study adherence and performance on M2C2 tasks, adjusting for covariates. All descriptive and regression analyses were conducted using SPSS version 28.0.1.1 (IBM Corp). Linear mixed-effects models were conducted using the *nlme* R package.

Ethical Considerations

The project received approval from the Butler Hospital Institutional Review Board (#1882523), and all the participants provided informed consent. All data was de-identified. Participants were compensated for their time with a $20 gift card.

Results

Participant Characteristics and Correlates of Subjective Cognitive Concerns

Overall sample demographics and SCC, adherence, satisfaction, motivation, and performance variables are presented in Table 1. The sample (N=122) was 68.85 (SD 4.93) years old

(range=60 - 81 y) with 16.52 (SD 2.39) years of formal education. The majority of the sample identified as female (n=82, 66.7%), White (n=106, 86.2%), and non-Hispanic (n=109, 88.6%). In the subset of individuals with Aβ PET scans, there were 25 Aβ positive participants and 48 Aβ negative participants. Subjective cognitive concerns, as measured on the CFI, were fairly minimal in this sample. On average, the sample scored 1.8 (SD 1.7) out of 14 possible points on the overall CFI ; however, the most common score on the CFI was 0 (mode=0), with 18% of the sample expressing no concerns about their cognition on the CFI.

Table . Sample demographic, subjective cognitive concerns, and mobile app protocol metrics.

	Total sample (N=122)	Subsample with Aβ[a] PET[b] (n=73)
Demographic characteristics		
Age (years), mean (SD)	68.85 (4.93)	69.25 (4.48)
Education (years), mean (SD)	16.52 (2.39)	16.45 (2.60)
Female, n (%)	82 (66.7)	52 (71.2)
White, n (%)	106 (86.2)	65 (89)
Non-Hispanic, n (%)	109 (88.6)	64 (87.7)
Depression (GDS[c]), mean (SD)	0.27 (0.54)	0.24 (0.46)
Loneliness (UCLA[d]), mean (SD)	0.70 (1.15)	0.56 (0.98)
Cognitive status (TICSm[e]), mean (SD)	39.32 (3.33)	39.01 (3.03)
Subjective cognitive concerns		
CFI[f] total score, mean (SD)	1.86 (1.70)	1.96 (1.80)
CFI-Cog[g], mean (SD)	1.39 (1.35)	1.96 (1.80)
Protocol adherence, mean (SD)	22.45 (2.58)	21.86 (2.30)
Protocol satisfaction		
Who had fun, n (%)	102 (85)	59 (83.1)
Who became bored, n (%)	30 (24.8)	17 (23.6)
Protocol motivation		
Who were motivated, n (%)	93 (77.5)	53 (73.6)
M2C2[h] task performance (average over 8-day protocol)		
Prices, mean (SD)	0.75 (0.09)	0.75 (0.09)
Color Shapes, mean (SD)	2.45 (0.76)	2.46 (0.80)
Symbol Search, mean (SD)	2224.78 (447.76)	2215.08 (410.92)

[a]Aβ: amyloid.

[b]PET: positron emission tomography.

[c]GDS: Geriatric Depression Scale.

[d]UCLA: University of California, Los Angeles.

[e]TICSm: modified Telephone Interview for Cognitive Status.

[f]CFI: Cognitive Function Instrument.

[g]CFI-Cog: Cognitive Function Instrument cognitive items.

[h]M2C2: Mobile Monitoring of Cognitive Change.

Bivariate associations are presented in Table 2. The CFI did show a small to moderate correlation with worse cognitive status (TICSm, r=−0.202, P=.03), greater loneliness (UCLA Loneliness Scale, r=0.248, P=.006), and worse adherence (r=−0.207, P=.02). In terms of overall M2C2 performances, the CFI showed a small correlation with episodic memory (ie, Prices) performance (r=−0.196, P=.03).

Table . Bivariate associations among demographic variables, subjective cognitive concerns, mobile app protocol metrics, and amyloid (Aβ) positron emission tomography (PET)[a].

		Age	Sex	Education	Depression	Loneliness	TICSm[b]	CFI[c]	CFI-Cog[d]	Adherence	Prices	Color Shape	Symbol Search	Aβ positive
Age														
	r	—[e]	−0.143	−0.005	−0.154	−0.024	−0.262	0.001	−0.020	−0.060	−0.369	−0.308	0.418	0.258
	P value	—	.12	.95	.09	.79	.004	.99	.83	.51	<.001	<.001	<.001	.03
Sex														
	r	−0.143	—	−0.066	−0.038	0.146	0.184	0.001	−0.050	0.031	0.221	0.271	−0.093	0.012
	P value	.12	—	.47	.68	.11	.04	.99	.60	.73	.01	.003	.31	.92
Education														
	r	−0.005	−0.066	—	0.011	−0.085	0.218	0.074	0.083	0.036	−0.028	0.020	−0.140	−0.015
	P value	.95	.47	—	.91	.36	.02	.42	.37	.70	.76	.83	.13	.90
Depression														
	r	−0.154	−0.038	0.011	—	0.170	0.063	0.040	0.064	0.054	0.202	−0.133	−0.028	0.052
	P value	.09	.68	.91	—	.06	.50	.67	.49	.56	.03	.15	.76	.66
Loneliness														
	r	−0.024	0.146	−0.085	0.170	—	−0.056	0.248	0.104	−0.094	0.001	0.033	−0.069	−0.031
	P value	.79	.11	.36	.06	—	.55	.006	.26	.31	.99	.72	.45	.80
TICSm														
	r	−0.262	0.184	0.218	0.063	−0.056	—	−0.202	−0.183	0.008	0.350	0.225	−0.216	−0.175
	P value	.004	.04	.02	.50	.55	—	.03	.05	.94	<.001	.01	.02	.14
CFI														
	r	0.001	0.001	0.074	0.040	0.248	−0.202	—	0.941	−0.176	−0.196	−0.028	−0.023	−0.002
	P value	.99	.99	.42	.67	.006	.03	—	<.001	.05	.03	.76	.80	.98
CFI-Cog														
	r	−0.020	−0.050	0.083	0.064	0.104	−0.183	0.941	—	−0.207	−0.171	−0.017	−0.038	−0.041
	P value	.83	.60	.37	.49	.26	.05	<.001	—	.02	.06	.86	.68	.73
Adherence														
	r	−0.060	0.031	0.036	0.054	−0.094	0.008	−0.176	−0.207	—	0.221	0.043	−0.075	0.210
	P value	.51	.73	.70	.56	.31	.94	.05	.02	—	.02	.64	.42	.08
Prices														
	r	−0.369	0.221	−0.028	0.202	0.001	0.350	−0.196	−0.171	0.221	—	0.309	−0.139	−0.325
	P value	<.001	.01	.76	.03	.99	<.001	.03	.06	.02	—	<.001	.13	.005
Color Shapes														
	r	−0.308	0.271	0.020	−0.133	0.033	0.225	−0.028	−0.017	0.043	0.309	—	−0.287	−0.141
	P value	<.001	.003	.83	.15	.72	.01	.76	.86	.64	<.001	—	.001	.24
Symbol Search														
	r	0.418	−0.093	−0.140	−0.028	−0.069	−0.216	−0.023	−0.038	−0.075	−0.139	−0.287	—	0.011
	P value	<.001	.31	.13	.76	.45	.02	.80	.68	.42	.13	.001	—	.93
Aβ positive														
	r	0.258	0.012	−0.015	0.052	−0.031	−0.175	−0.002	−0.041	0.210	−0.325	−0.141	0.011	—
	P value	.03	.92	.90	.66	.80	.14	.98	.73	.08	.005	.24	.93	—

[a]The Aβ PET data was only available for a subsample (n=73).

[b]TICSm: modified Telephone Interview for Cognitive Status

[c]CFI: Cognitive Function Instrument

[d]CFI-Cog: Cognitive Function Instrument cognitive items

[e]Not applicable

Aim 1: Subjective Cognitive Concerns and Adherence to Mobile App–Based Cognitive Assessment Protocol

Overall, remote assessment protocol adherence was high in the sample, with participants completing on average 93.5% of the 24 testing sessions over the 8-day period. CFI-Cog endorsement significantly predicted overall remote testing adherence over 8 days. This suggests that higher levels of SCC were associated with worse adherence (a diminished number of test sessions completed across the protocol) (ß=−.197, *P*=.04, 95% CI −.647 to −.021). However, this association was not observed when using the CFI total score as the predictor (Table 3).

Table . Association between subjective cognitive concerns and mobile app metrics.

	R^2	ß/aOR[a] (95% CI)	*P* value
Aim 1: Subjective cognitive concerns and adherence to app-based cognitive assessment protocol, ß			
Adherence	0.042		
SCC[b] (CFI[c])		−.157 (−.459 to .042)	.10
Age		−.046 (−.106 to .064)	.63
Sex		.051 (−.657 to 1.150)	.59
Depression (GDS[d])		.072 (−.479 to 1.068)	.45
Loneliness (UCLA[e])		−.071 (−.5.20 to .239)	.47
Aim 2: Subjective cognitive concerns, satisfaction, and motivation on app-based cognitive assessment, aOR			
Protocol satisfaction			
Level of fun	0.060		
SCC (CFI)		1.016 (0.738 to 1.398)	.92
Age		0.938 (0.844 to 1.042)	.23
Sex		0.328 (0.112 to 0.963)	.04
Depression (GDS)		0.679 (0.28 to 1.606)	.38
Loneliness (UCLA)		0.883 (0.554 to 1.408)	.60
Boredom	0.026		
SCC (CFI)		1.013 (0.783 to 1.310)	.92
Age		1.056 (0.966 to 1.155)	.23
Sex		0.575 (0.213 to 1.552)	.27
Depression (GDS)		1.060 (0.473 to 2.375)	.89
Loneliness (UCLA)		1.130 (0.785 to 1.625)	.12
Protocol motivation			
Motivation	0.028		
SCC (CFI)		0.891 (0.691 to 1.148)	.37
Age		0.962 (0.877 to 1.054)	.41
Sex		1.689 (0.613 to 4.603)	.31
Depression (GDS)		0.685 (0.312 to 1.503)	.35
Loneliness (UCLA)		1.286 (0.831 to 1.991)	.26
Aim 3: Subjective cognition and performance on app based cognitive assessment, ß			
Prices	0.234		
SCC (CFI)		−.200 (−.020 to −.002)	.02
Age		−.318 (−.009 to −.003)	<.001
Sex		.193 (.005 to .070)	.02
Depression (GDS)		.174 (.001 to .056)	.04
Loneliness (UCLA)		−.013 (−.015 to .013)	.88
Color Shapes	0.170		
SCC (CFI)		−.016 (−.085 to .071)	.86

	R^2	ß/aOR[a] (95% CI)	*P* value
Age		−.301 (−.073 to −.020)	<.001
Sex		.205 (.051 to .612)	.02
Depression (GDS)		−.181 (−.489 to −.009)	.04
Loneliness (UCLA)		.030 (−.098 to .138)	.74
Symbol Search	0.178		
SCC (CFI)		−.007 (−48.34 to 44.72)	.94
Age		.416 (22.302 to 53.977)	<.001
Sex		−.027 (−192.599 to 140.950)	.76
Depression (GDS)		.041 (−109.610 to 176.575)	.64
Loneliness (UCLA)		−.045 −92.465 to 55.043	.62

[a]aOR: adjusted odds ratio.

[b]SCC: subjective cognitive concern.

[c]CFI: Cognitive Function Instrument.

[d]GDS: Geriatric Depression Scale.

[e]UCLA: University of California, Los Angeles.

Aim 2: Subjective Cognitive Concerns, Satisfaction, and Motivation on Mobile App–Based Cognitive Assessment

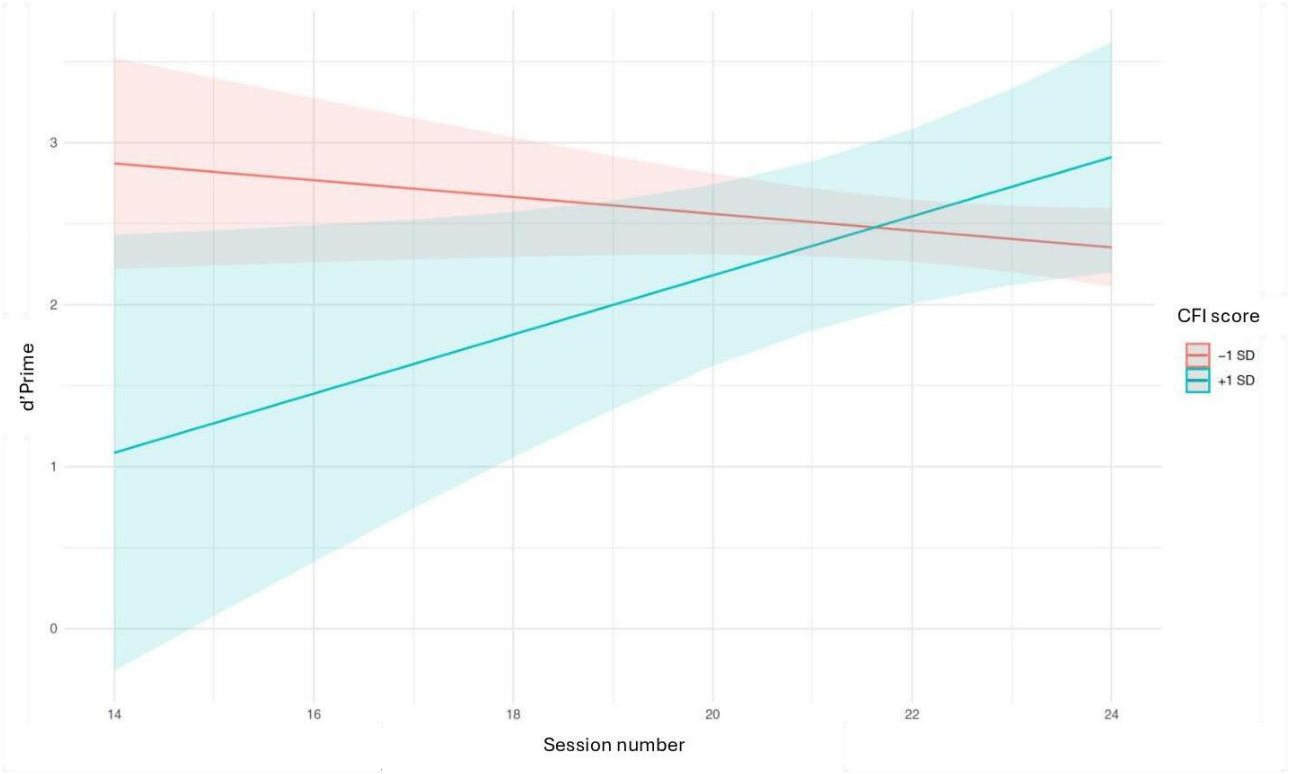
In terms of overall protocol satisfaction, participants endorsed that completing the assessments was fun (n=102, 85%) and not boring (n=92, 75.2%). Participants also reported that they did not encounter difficulties in motivating themselves to give their best performance (n=93, 77.5%). CFI total score was not associated with metrics of study satisfaction in the adjusted models, including how fun or boring participants found the protocol to be. In terms of motivation, CFI total score was not predictive of self-reported effort across the protocol period (more details in Table 3).

Aim 3: Subjective Cognitive Concerns and Performance on Mobile App–Based Cognitive Assessment

SCCs and their association with overall objective cognitive performances on digital cognitive tests was also investigated (more details in Table 3). High SCC endorsement on the CFI total score was significantly predictive of poorer overall performance on the Prices task (ß=−.200, *P*=.02, 95% CI −.020, to −.002). There were no significant associations between the CFI and performance metrics for Color Shapes (*P*=.86) or Symbol Search (*P*=.94).

We also investigated the association of SCC and longitudinal trends of cognitive performance over the 8-day testing period (M2C2 app performance descriptives for each study day have been previously reported) [38]. There was a main effect of the CFI total score suggesting that higher levels of SCC were associated with worse performance on Color Shapes (working memory) at day 1 (unstandardized estimate [b]=−1.047, SE=0.47, *P*=.03). Furthermore, there was a significant interaction between SCCs and Color Shapes performance (b=0.048, SE=0.02, *P*=.03), such that higher CFI total scores were associated with improved performance on this subtest over time. In other words, although high SCC was associated with worse working memory performance initially, individuals seemed to benefit from repeated practice and ultimately performed better (more details in Figure 3). There were no significant associations between SCC and intraindividual variability in Prices or Symbol Match performances over the protocol period.

Figure 3. Subjective cognitive concerns association with working memory longitudinally over the 8-day mobile-appmobile app protocoll. CFI: Cognitive Function Instrument.



Secondary Aim: Moderating Effect of Amyloid Status on the Relationship Between Subjective Cognitive Concerns and Mobile App–Based Cognitive Assessment

The moderation of Aβ positivity on the relationship between CFI and app-based metrics, as well as CFI and cognitive performance (ie, Prices, Color Shapes, and Symbol Search) was investigated in a subset of individuals with Aβ PET scans (n=73). The presence of Aβ did not moderate the relationship between CFI total score and adherence. There was evidence for moderation at the $P<.10$ level of Aβ positivity on the relationship between CFI total score and Prices was observed, but did not fall within the $\alpha=.05$ threshold (ß=.179, $P=.09$, 95% CI −.003 to .036). Aβ status did not moderate the association between SCC and Color Shapes or Symbol Search.

Discussion

Principal Findings

This study sought to characterize the relationship between SCC and mobile app–based protocol engagement (adherence, satisfaction, and motivation) and performance on cognitive tasks in a cohort of cognitively healthy older adults. While protocol satisfaction and motivation were not impacted by SCCs, overall adherence was associated, such that higher levels of SCCs predicted lower rates of adherence. Overall, we show that SCCs were associated with worse objective performance on mobile app–based assessments. Our findings suggest that mobile app–based assessments hold promise as sensitive tools to detect subtle neurodegenerative changes in individuals at risk for dementia, with potential application for early detection, diagnosis, and treatment.

We observed an association between SCCs, as measured on the Cognitive Function Index (CFI), and overall worse performance on a test of episodic memory (Prices). A significant association was not observed between SCC and overall performance on tasks of processing speed or working memory. Findings in the broader SCD literature are quite mixed as to whether SCCs are or are not associated with concurrent cognitive functioning abilities [52]; as such, our results may indicate that the use of sensitive digital cognitive tools could enhance the association between SCC and objective cognitive metrics. Our result reflecting a unique association with episodic memory is perhaps clinically meaningful, as memory concerns specifically are frequently considered to be the hallmark presenting symptom of AD [2,53], particularly in individuals over the age of 65 [53]. This observation is underscored by our finding of a moderating effect of cerebral Aβ positivity (at the $P<.10$ level) on the relationship between SCC and episodic memory (Prices) in a subsample of participants with biomarkers available for analysis (secondary aim), such that the association of SCC with episodic memory was stronger in the presence of Aβ. This result adds to the growing evidence base for an association between SCC and higher Aβ [52]. Replication of this analysis in a larger sample may clarify the interplay between SCC, AD biomarkers, and subtle memory deficits observed on mobile app–based assessment in otherwise cognitively and clinically normal older adults. That being said, it is worth considering that the unique association between SCC and episodic memory in our sample may be at least somewhat secondary to the way the CFI is constructed, as this instrument mostly focuses on episodic memory-related cognitive concerns [45]. Previous studies have

demonstrated the importance of querying beyond just memory concerns to most sensitively detect dementia risk [54,55]. Indeed, SCC can be sampled by a multitude of techniques and approaches (eg, memory only vs multiple cognitive domains; traditional paper and pencil vs digital assessments; self vs informant report; capturing current ability vs change in ability, etc) [24,25], all of which may impact the sensitivity of the SCC measures to detect current and risk for future cognitive impairment. Future research should carefully weigh SCC assessment approaches, and consider incorporating a more comprehensive SCC screener which queries across a broader range of cognitive domains, as this may be valuable for detecting more meaningful associations between SCC and mobile app–based cognitive performances.

We also evaluated patterns of cognitive performance across the 8-day assessment period. Analyses revealed that more SCC was associated with initially worse working memory (Color Shapes) performance on day one. However, individuals with higher SCC showed progressively improving working memory performance over the assessment period, and by day 8 were performing in a similar range with those with lower SCC. This pattern was not observed on the episodic memory or processing speed subtests. This finding is consistent with recent results from Aschenbrenner et al [56], which showed variability over high frequency app-based assessments in attention and working memory-based skills, but not in episodic memory, in older adults who were at increased genetic risk for AD [56]. Our finding of an initial relative weakness in working memory skills could reflect the presence of mild cognitive inefficiencies in older adults with SCC, which may be compensated for with repeated exposure to the task over time. This may be due to early dysfunction in working memory processes, which have been characterized in preclinical AD [57,58]. Subjectively, this may be experienced by these individuals as thinking which requires increased concentration, that is punctuated by occasional lapses, and is associated perhaps with a sense of being cognitively overwhelmed, all of which may be driving the report SCC in daily life, but may be subtle enough that it may not be detected on traditional neuropsychological tests. Hence, this type of unique marker of cognitive inefficiency would only be afforded through high frequency, mobile app–based cognitive assessment in this population. The plausible role of performance anxiety may also explain variability in cognitive performance overtime. Although beyond the scope of the present analysis, research which comprehensively measures the role of in-the-moment psychological and physiological states by ecological momentary assessment techniques in relation to in the moment SCC and objective cognition will be critical next steps.

Our findings related to SCC and metrics of protocol engagement and satisfaction with unsupervised, high frequency digital cognitive assessments have potentially broad implications for aging research, especially in light of increasing interest in using these techniques to remotely assess cognitive functioning in this population [31,32]. Indeed, whether or not a study is directly examining SCC or includes SCD as a clinical group of interest, aging studies must contend with the fact that SCCs are widespread and may be clinically meaningful feature among the older adults in their sample [2]. We show that SCC was not associated with protocol satisfaction (including how fun or how boring the experience was) nor did SCC effect how motivated a participant was to engage and put forth their best effort. However, higher levels of SCC did impact protocol adherence in our sample, with mildly worse compliance those with higher SCC observed. For context, protocol adherence was overall quite high in our population (93.5%) and similarly high compliance rates have been described in several other recent smartphone-based studies [32], yet no previous studies have directly explored the impact of SCCs on protocol adherence. It may be that high frequency, mobile assessments are able to detect mild patterns of forgetfulness (manifested in slightly worse adherence rates) in older adults with SCC, so future research should be attuned to this possibility. Replication of this association between SCC and adherence to remote protocols in larger, more demographically diverse samples is warranted.

Limitations

Some limitations of this study, as well as potential future directions of this research, warrant acknowledgment. Our sample was largely White, disproportionally female, and highly educated. This limits the generalizability of the study findings to broader demographic groups. Our sample was from a relatively small single cohort and only a subset had AD biomarkers available for analysis, restricting study power and may have impeded our ability to obtain significant findings. A measure of anxiety symptomology was not available in the present sample, limiting our ability to understand the plausible role of anxiety, SCC, and performance on digital assessments [32,59]. The main predictor of interest in this study, the CFI, shows adequate evidence for validity and reliability [42], but is somewhat limited in the domains of cognitive concerns it measures (eg, memory and daily functioning). Research has shown that comprehensive assessment of concerns across a broad range of cognitive domains may be most sensitive to risk in aging populations [60], so future smartphone-based research may consider including a broader range of cognitive concern items to better understand this relationship. We were unable to assess informant-report of SCC in this sample, which has been demonstrated to be optimized to predict risk above and beyond self-report [25], and this should be addressed in future smartphone studies. Although such metrics were not available in this study, there is also a small but growing literature which has examined the use of smartphones to assess SCC in a digital remote format in tandem with cognitive tests and relevant clinical outcomes in older adult samples [61-63]. Future research on SCC and smartphone-based digital cognitive assessments should investigate the relative value of traditional in person reported SCC (eg, the CFI) versus smartphone based digital SCC assessments. Finally, this cross-sectional study was not able to investigate the relationship between SCC and longitudinal decline on smartphone-based cognitive performance in individuals with positive AD biomarkers. It will be essential for future studies to follow participants over time to understand if SCC at baseline may predict incident decline on these novel cognitive metrics. Due to the limited nature of this pilot study, we cannot provide specific recommendations about optimal SCC assessment approach for use in future digital cognitive studies.

Conclusions

SCCs are frequently one of the earliest clinical symptoms of dementia [2]. However, clinical interpretations of SCCs are complicated by the ubiquitous endorsement of these concerns [2], their potential for non-neurodegenerative etiology (eg, psychiatric, medical, or sociodemographic factors) [2,3], and frequent lack of association with traditional objective cognitive testing [1]. Smartphone-based, digital cognitive assessments, which are increasingly used in aging research, may offer potential for improved ecological validity and sensitivity to subtle markers of cognitive decline [30-32,35]. Hence, such smartphone-based digital tools may be able to capture quite mild deficits in individuals with SCC who perform within normal limits on traditional neuropsychological tests. Results from the current study showed that SCCs are associated with worse overall memory performance on mobile app assessment, and patterns of cognitive inefficiency (variable working memory) and mild forgetfulness (diminished adherence) across an 8-day assessment period in cognitively intact older adults. Findings indicate that mobile app assessments may be uniquely sensitive to very subtle neurodegenerative changes in at risk older adults, with critical implications for early detection and timely intervention.

Acknowledgments

This work is supported by Alzheimer's Association grant AACSF-20-685786 (Thompson, PI) and by NIA grants T32 AG049676 and U2CAG060408 to Pennsylvania State University.

Authors' Contributions

CON took the lead on study conceptualization, formal analysis, original draft writing, and visualization. ANDV and ZJK contributed to formal analyses, visualization, and draft writing and editing. SP played an instrumental role in draft writing and editing. JS, KDH, NR, MJS, and LAR provided valuable feedback on study conceptualization and draft review and editing. LIT was responsible for funding acquisition, project administration, study conceptualization, and draft writing, review, and editing.

Conflicts of Interest

KDH is an employee of Cogstate Ltd. All remaining authors declare no conflicts of interest.

References

1. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alz Dement 2014 Nov;10(6):844-852. [doi: 10.1016/j.jalz.2014.01.001] [Medline: 24798886]
2. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. Lancet Neurol 2020 Mar;19(3):271-278. [doi: 10.1016/S1474-4422(19)30368-0] [Medline: 31958406]
3. Rabin LA, Smart CM, Amariglio RE. Subjective cognitive decline in preclinical Alzheimer's disease. Annu Rev Clin Psychol 2017 May 8;13:369-396. [doi: 10.1146/annurev-clinpsy-032816-045136] [Medline: 28482688]
4. Wang XT, Wang ZT, Hu HY, et al. Association of subjective cognitive decline with risk of cognitive impairment and dementia: a systematic review and meta-analysis of prospective longitudinal studies. J Prev Alzheimers Dis 2021;8(3):277-285. [doi: 10.14283/jpad.2021.27] [Medline: 34101784]
5. Parfenov VA, Zakharov VV, Kabaeva AR, Vakhnina NV. Subjective cognitive decline as a predictor of future cognitive decline: a systematic review. Dem Neuropsychol 2020;14(3):248-257. [doi: 10.1590/1980-57642020dn14-030007] [Medline: 32973979]
6. Pike KE, Cavuoto MG, Li L, Wright BJ, Kinsella GJ. Subjective cognitive decline: level of risk for future dementia and mild cognitive impairment, a meta-analysis of longitudinal studies. Neuropsychol Rev 2022 Dec;32(4):703-735. [doi: 10.1007/s11065-021-09522-3] [Medline: 34748154]
7. Valech N, Sánchez-Benavides G, Tort-Merino A, et al. Associations between the subjective cognitive decline-questionnaire's scores, gray matter volume, and amyloid-β levels. J Alzheimers Dis 2019;72(4):1287-1302. [doi: 10.3233/JAD-190624]
8. Ryu SY, Lim EY, Na S, et al. Hippocampal and entorhinal structures in subjective memory impairment: a combined MRI volumetric and DTI study. Int Psychogeriatr 2017 May;29(5):785-792. [doi: 10.1017/S1041610216002349] [Medline: 28067183]
9. Yue L, Wang T, Wang J, et al. Asymmetry of hippocampus and amygdala defect in subjective cognitive decline among the community dwelling Chinese. Front Psychiatry 2018;9:226. [doi: 10.3389/fpsyt.2018.00226] [Medline: 29942265]
10. Yan T, Wang W, Yang L, Chen K, Chen R, Han Y. Rich club disturbances of the human connectome from subjective cognitive decline to Alzheimer's disease. Theranostics 2018;8(12):3237-3255. [doi: 10.7150/thno.23772] [Medline: 29930726]
11. van Rooden S, van den Berg-Huysmans AA, Croll PH, et al. Subjective cognitive decline is associated with greater white matter hyperintensity volume. J Alzheimers Dis 2018;66(3):1283-1294. [doi: 10.3233/JAD-180285] [Medline: 30412485]
12. Sacuiu S, Eckerström M, Johansson L, et al. Increased risk of dementia in subjective cognitive decline if CT brain changes are present. J Alzheimers Dis 2018;66(2):483-495. [doi: 10.3233/JAD-180073]

13. Ohlhauser L, Parker AF, Smart CM, Gawryluk JR. White matter and its relationship with cognition in subjective cognitive decline. Alz Dem Diag Ass Dis Mo 2019 Dec;11(1):28-35. [doi: 10.1016/j.dadm.2018.10.008]
14. Chimthanawala NMA, Haria A, Sathaye S. Non-invasive biomarkers for early detection of Alzheimer's disease: a new-age perspective. Mol Neurobiol 2024 Jan;61(1):212-223. [doi: 10.1007/s12035-023-03578-3] [Medline: 37596437]
15. Rostamzadeh A, Bohr L, Wagner M, Baethge C, Jessen F. Progression of subjective cognitive decline to MCI or Dementia in relation to biomarkers for Alzheimer disease. Neurology (ECronicon) 2022 Oct 25;99(17):e1866-e1874. [doi: 10.1212/WNL.0000000000201072]
16. Wolfsgruber S, Molinuevo JL, Wagner M, et al. Prevalence of abnormal Alzheimer's disease biomarkers in patients with subjective cognitive decline: cross-sectional comparison of three European memory clinic samples. Alzheimers Res Ther 2019 Jan 17;11(1):8. [doi: 10.1186/s13195-018-0463-y] [Medline: 30654834]
17. Jadick MF, Robinson T, Farrell ME, et al. Associations between self and study partner report of cognitive decline with regional tau in a multicohort study. Neurology (ECronicon) 2024 Jun 25;102(12):e209447. [doi: 10.1212/WNL.0000000000209447] [Medline: 38810211]
18. Hao L, Sun Y, Li Y, et al. Demographic characteristics and neuropsychological assessments of subjective cognitive decline (SCD) (plus). Ann Clin Transl Neurol 2020 Jun;7(6):1002-1012. [doi: 10.1002/acn3.51068] [Medline: 32588989]
19. Jenkins A, Tree JJ, Thornton IM, Tales A. Subjective cognitive impairment in 55-65-year-old adults is associated with negative affective symptoms, neuroticism, and poor quality of life. J Alzheimer's Dis 2019;67(4):1367-1378. [doi: 10.3233/JAD-180810]
20. Kielb S, Rogalski E, Weintraub S, Rademaker A. Objective features of subjective cognitive decline in a United States national database. Alzheimers Dement 2017 Dec;13(12):1337-1344. [doi: 10.1016/j.jalz.2017.04.008]
21. Wolfsgruber S, Kleineidam L, Guski J, et al. Minor neuropsychological deficits in patients with subjective cognitive decline. Neurology (ECronicon) 2020 Sep 1;95(9):e1134-e1143. [doi: 10.1212/WNL.0000000000010142] [Medline: 32636322]
22. Webster-Cordero F, Giménez-Llort L. The challenge of subjective cognitive complaints and executive functions in middle-aged adults as a preclinical stage of dementia: a systematic review. Geriatrics (Basel) 2022 Mar 8;7(2):30. [doi: 10.3390/geriatrics7020030] [Medline: 35314602]
23. Stark M, Wolfsgruber S, Kleineidam L, et al. Minor neuropsychological deficits in SCD: association with AD biomarkers and cognitive decline. Alz Dement 2023 Dec;19(S18):e073332. [doi: 10.1002/alz.073332]
24. Rabin LA, Smart CM, Crane PK, et al. Subjective cognitive decline in older adults: an overview of self-report measures used across 19 international research studies. J Alzheimers Dis 2015 Sep 24;48:S63-S86. [doi: 10.3233/JAD-150154] [Medline: 26402085]
25. Chang KH, Wang C, Nester CO, et al. Examining the role of participant and study partner report in widely-used classification approaches of mild cognitive impairment in demographically-diverse community dwelling individuals: results from the Einstein Aging Study. Front Aging Neurosci 2023;15:1221768. [doi: 10.3389/fnagi.2023.1221768] [Medline: 38076542]
26. Butterbrod E, Rabin L, Tommet D, et al. Toward optimizing measurement of self‐perceived cognitive function‐a global expert survey on behalf of the SCD‐I Working Group. Alz Dement 2023 Dec;19(S18):e078216. [doi: 10.1002/alz.078216]
27. Rabin LA, Sikkes SAM, Tommet D, et al. Linking self-perceived cognitive functioning questionnaires using item response theory: the subjective cognitive decline initiative. Neuropsychology 2023 May;37(4):463-499. [doi: 10.1037/neu0000888] [Medline: 37276136]
28. Jessen F. Subjective and objective cognitive decline at the pre-dementia stage of Alzheimer's disease. Eur Arch Psychiatry Clin Neurosci 2014 Nov;264 Suppl 1:S3-S7. [doi: 10.1007/s00406-014-0539-z] [Medline: 25238934]
29. Gelles-Watnick R. Americans' use of mobile technology and home broadband. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/2024/01/31/americans-use-of-mobile-technology-and-home-broadband [accessed 2025-01-15]
30. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
31. Öhman F, Hassenstab J, Berron D, Schöll M, Papp KV. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alz Dem Diag Ass Dis Mo 2021 Jan;13(1):e12217. [doi: 10.1002/dad2.12217]
32. Harris C, Tang Y, Birnbaum E, Cherian C, Mendhe D, Chen MH. Digital neuropsychology beyond computerized cognitive assessment: applications of novel digital technologies. Arch Clin Neuropsychol 2024 Apr 24;39(3):290-304. [doi: 10.1093/arclin/acae016] [Medline: 38520381]
33. Leroy V, Gana W, Aïdoud A, et al. Digital health technologies and Alzheimer's disease clinical trials: might decentralized clinical trials increase participation by people with cognitive impairment? Alz Res Therapy 2023;15(1):87. [doi: 10.1186/s13195-023-01227-4]
34. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2018 Jan;25(1):14-30. [doi: 10.1177/1073191116643164] [Medline: 27084835]
35. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multiday Boston Remote Assessment of Neurocognitive Health (BRANCH): feasibility, reliability, and validity. Neuropsychology 2024 Feb;38(2):198-210. [doi: 10.1037/neu0000933] [Medline: 37971862]

36. Cerino ES, Katz MJ, Wang C, et al. Variability in cognitive performance on mobile devices is sensitive to mild cognitive impairment: results from the Einstein Aging Study. Front Digit Health 2021;3:758031. [doi: 10.3389/fdgth.2021.758031] [Medline: 34927132]
37. Benge JF, Kiselica AM, Aguirre A, et al. Technology use and subjective cognitive concerns in older adults. Arch Gerontol Geriatr 2023 Mar;106:104877. [doi: 10.1016/j.archger.2022.104877] [Medline: 36459914]
38. Thompson LI, Vito AN, Kunicki Z, et al. The feasibility and psychometric rigor of using high-frequency, smartphone-based assessments for cognitive screening in older adults. PsychArchives 2024. [doi: 10.1017/S1355617724000328]
39. Thompson LI, Cummings M, Emrani S, et al. Digital clock drawing as an Alzheimer's disease susceptibility biomarker: associations with genetic risk score and APOE in older adults. J Prev Alzheimers Dis 2024;11(1):79-87. [doi: 10.14283/jpad.2023.48] [Medline: 38230720]
40. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. Ann Neurol 2024 Mar;95(3):507-517. [doi: 10.1002/ana.26833] [Medline: 37991080]
41. Thompson LI, Kunicki ZJ, Emrani S, et al. Remote and in‐clinic digital cognitive screening tools outperform the MoCA to distinguish cerebral amyloid status among cognitively healthy older adults. Alz Dem Diag Ass Dis Mo 2023 Oct;15(4):e12500. [doi: 10.1002/dad2.12500]
42. Amariglio RE, Sikkes SAM, Marshall GA, et al. Item-level investigation of participant and study partner report on the Cognitive Function Index from the A4 study screening data. J Prev Alzheimers Dis 2021;8(3):257-262. [doi: 10.14283/jpad.2021.8] [Medline: 34101781]
43. Alber J, Lee AKW, Menard W, Monast D, Salloway SP. Recruitment of at-risk participants for clinical trials: a major paradigm shift for Alzheimer's disease prevention. J Prev Alzheimers Dis 2017;4(4):213-214. [doi: 10.14283/jpad.2017.32] [Medline: 29181484]
44. Thompson LI, Harrington KD, Roque N, et al. A highly feasible, reliable, and fully remote protocol for mobile app‐based cognitive assessment in cognitively healthy older adults. Alz Dem Diag Ass Dis Mo 2022 Jan;14(1):e12283. [doi: 10.1002/dad2.12283]
45. Li C, Neugroschl J, Luo X, et al. The utility of the Cognitive Function Instrument (CFI) to detect cognitive decline in non-demented older adults. J Alzheimer's Dis 2017;60(2):427-437. [doi: 10.3233/JAD-161294]
46. Rinaldi P, Mecocci P, Benedetti C, et al. Validation of the five-item geriatric depression scale in elderly subjects in three different settings. J Am Geriatr Soc 2003 May;51(5):694-698. [doi: 10.1034/j.1600-0579.2003.00216.x] [Medline: 12752847]
47. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
48. Cook SE, Marsiske M, McCoy KJM. The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnestic mild cognitive impairment. J Geriatr Psychiatry Neurol 2009 Jun;22(2):103-109. [doi: 10.1177/0891988708328214] [Medline: 19417219]
49. Robinson T, Klinger H, Buckley R, et al. Subjective cognitive decline across ethnoracial groups in the A4 study. Alzheimers Dement 2023 Sep;19(9):4084-4093. [doi: 10.1002/alz.13138]
50. Li C, Hong Y, Yang X, et al. The use of subjective cognitive complaints for detecting mild cognitive impairment in older adults across cultural and linguistic groups: a comparison of the Cognitive Function Instrument to the Montreal Cognitive Assessment. Alz Dement 2023 May;19(5):1764-1774. [doi: 10.1002/alz.12804]
51. Hakun J, Elbich DB, Roque NA, Yabiku ST, Sliwinski M. Mobile monitoring of cognitive change (M2C2): high-frequency assessments and protocol reporting guidelines. PsyArXiv. Preprint posted online on Aug 22, 2024. [doi: 10.31234/osf.io/34ux5]
52. Munro CE, Boyle R, Chen X, et al. Recent contributions to the field of subjective cognitive decline in aging: a literature review. Alz Dem Diag Ass Dis Mo 2023 Oct;15(4):e12475. [doi: 10.1002/dad2.12475]
53. Barnes J, Dickerson BC, Frost C, Jiskoot LC, Wolk D, van der Flier WM. Alzheimer's disease first symptoms are age dependent: evidence from the NACC dataset. Alzheimers Dement 2015 Nov;11(11):1349-1357. [doi: 10.1016/j.jalz.2014.12.007]
54. La Joie R, Perrotin A, Egret S, et al. Qualitative and quantitative assessment of self‐reported cognitive difficulties in nondemented elders: association with medical help seeking, cognitive deficits, and β‐amyloid imaging. Alz Dem Diag Ass Dis Mo 2016 Jan;5(1):23-34. [doi: 10.1016/j.dadm.2016.12.005]
55. Nester C, Ayers E, Rabin L, Verghese J. Non-memory subjective cognitive concerns predict incident motoric cognitive risk syndrome. Eur J Neurol 2020 Jul;27(7):1146-1154. [doi: 10.1111/ene.14271] [Medline: 32319724]
56. Aschenbrenner AJ, Hassenstab J, Morris JC, Cruchaga C, Jackson JJ. Relationships between hourly cognitive variability and risk of Alzheimer's disease revealed with mixed-effects location scale models. Neuropsychology 2024 Jan;38(1):69-80. [doi: 10.1037/neu0000905] [Medline: 37079810]
57. Huntley JD, Howard RJ. Working memory in early Alzheimer's disease: a neuropsychological review. Int J Geriat Psychiatry 2010 Feb;25(2):121-132. [doi: 10.1002/gps.2314]
58. van Geldorp B, Heringa SM, van den Berg E, Olde Rikkert MGM, Biessels GJ, Kessels RPC. Working memory binding and episodic memory formation in aging, mild cognitive impairment, and Alzheimer's dementia. J Clin Exp Neuropsychol 2015;37(5):538-548. [doi: 10.1080/13803395.2015.1037722] [Medline: 26011711]

59. Liew TM. Subjective cognitive decline, anxiety symptoms, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Sep 11;12(1):107. [doi: 10.1186/s13195-020-00673-8] [Medline: 32917264]
60. Nester CO, Gao Q, Katz MJ, et al. Does the cognitive change index predict future cognitive and clinical decline? Longitudinal analysis in a demographically diverse cohort. J Alzheimers Dis 2024;98(1):319-332. [doi: 10.3233/JAD-230752] [Medline: 38393900]
61. Rabin L, Mogle J, Katz MJ, Sliwinski MJ. Do repeated ambulatory cognitive assessments enhance the relation between objective and subjective cognition in non‐demented community dwelling older adults? Results from the Einstein Aging Study. Alzheimers Dement 2020 Dec;16(S6):e045918. [doi: 10.1002/alz.045918]
62. Mogle J, Turner JR, Bhargava S, Stawski RS, Almeida DM, Hill NL. Individual differences in frequency and impact of daily memory lapses: results from a national lifespan sample. BMC Geriatr 2023 Oct 17;23(1):670. [doi: 10.1186/s12877-023-04363-6] [Medline: 37848825]
63. Rabin L, Mogle J, Wang C, et al. Cognitive lapses reported by daily digital diary, and not by recall‐based questionnaires, predict future MCI in a demographically diverse cohort of community‐dwelling older adults. Alzheimers Dement 2023 Dec;19(S22):e077994. [doi: 10.1002/alz.077994]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
Aβ: amyloid
CFI: Cognitive Function Instrument
CFI-Cog: Cognitive Function Instrument cognitive items
GDS: Geriatric Depression Scale
M2C2: Mobile Monitoring of Cognitive Change
MAP: Memory and Aging Program
PET: positron emission tomography
SCC: subjective cognitive concern
SCD: subjective cognitive decline
TICSm: modified Telephone Interview for Cognitive Status
UCLA: University of California, Los Angeles

Edited by J Job; submitted 08.07.24; peer-reviewed by I Liampas, S Young; revised version received 24.10.24; accepted 19.11.24; published 04.02.25.

Please cite as:
Nester CO, De Vito AN, Prieto S, Kunicki ZJ, Strenger J, Harrington KD, Roque N, Sliwinski MJ, Rabin LA, Thompson LI
Association of Subjective Cognitive Concerns With Performance on Mobile App–Based Cognitive Assessment in Cognitively Normal Older Adults: Observational Study
JMIR Aging 2025;8:e64033
URL: https://aging.jmir.org/2025/1/e64033
doi:10.2196/64033

Original Paper

Effectiveness of a Dyadic Technology–Enhanced Home-Based Horticultural Therapy on Psychosocial Well-Being Among People With Dementia and Their Family Caregivers: Multimethods Pilot Study

Patrick Pui Kin Kor[1], RN, PhD; Justina Yat Wa Liu[1,2], RN, PhD; Arkers Kwan Ching Wong[1], RN, PhD; Alex Pak Lik Tsang[1], CPsychol, PhD; Han Zhi Tan[1], MSc; Daphne Sze Ki Cheung[1,3,4], RN, PhD; Humphrey Kwong Wai Leung[5], DBA, FHKEng; Frances Kam Yuet Wong[1], RN, PhD

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

[2]Research Institute for Smart Ageing, PolyU Academy for Interdisciplinary Research, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]School of Nursing and Midwifery, Faculty of Health, Deakin University, Melbourne, Australia

[4]Centre for Quality and Patient Safety Research/Alfred Health Partnership, Institute for Health Transformation, Deakin University, Melbourne, Australia

[5]University Fellows Association, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Patrick Pui Kin Kor, RN, PhD
School of Nursing
The Hong Kong Polytechnic University
11 Yuk Choi Road
Hung Hom, Kowloon
Hong Kong
China (Hong Kong)
Phone: 852 27665622
Email: patrick.kor@polyu.edu.hk

Abstract

Background: Horticultural therapy (HT) has been proposed to be an effective intervention for improving the psychosocial well-being of people with dementia and their caregivers. However, constraints such as limited land space in high-density cities, unstable weather, and lack of gardening experience may hamper the delivery of HT to people with dementia and their caregivers.

Objective: This pilot study aimed to examine the feasibility and preliminary effects of a technology-enhanced home-based HT for people with dementia and their caregivers using a hydroponic indoor growing system.

Methods: A single-group pre-post design was adopted. A total of 37 dyads of people with dementia and their caregivers participated in 3 weekly face-to-face sessions, followed by 8 weeks of home-based horticultural activities. Outcomes were measured at baseline and postintervention (at week 11), including feasibility outcomes, cognitive function, neuropsychiatric symptoms, and happiness levels of people with dementia. Caregivers' outcomes included positive aspects of caregiving, perceived stress levels, depressive symptoms, caregiver distress, and happiness levels. Semistructured focus group interviews were conducted with the caregivers to further explore their horticultural experience.

Results: Intervention feasibility was established with a completion rate of 83.78% and an attrition rate of 2.63% (n=1). Significant improvements were detected in caregiver distress (P<.05) and the happiness level of people with dementia (P<.01). The qualitative findings indicated that HT improved the psychological well-being of both people with dementia and caregivers, enhanced the relationships between caregivers and people with dementia, expanded the caregivers' social networks, and enhanced the autobiographical memory of people with dementia.

Conclusions: This pilot study provides evidence on the feasibility of using a hydroponic indoor grower to conduct home-based HT for people with dementia and their caregivers. The findings suggest positive effects on the psychological well-being of both people with dementia and their caregivers. Caregivers reported potential positive effects of HT on the autobiographical memory retrieval of people with dementia. Due to the pilot nature of this study, a control group was not employed. Therefore, large-scale randomized controlled trials are encouraged to further confirm the effectiveness of the intervention.

Trial Registration: ClinicalTrials.gov NCT05577975; https://clinicaltrials.gov/study/NCT05577975

(JMIR Aging 2025;8:e66017) doi:10.2196/66017

KEYWORDS

horticultural activity; dementia; caregivers; dyadic intervention; technology–enhanced intervention

Introduction

With the global aging population, dementia has become a significant public health concern. The World Health Organization estimates that the number of people with dementia worldwide is approximately 55 million, with this figure expected to rise to about 78 million by 2030 and 139 million by 2050 [1]. The global cost of dementia saw a substantial increase of 62% between 2010 and 2019, with the economic burden projected to reach an estimated US $1313.4 billion in 2019. Notably, informal care significantly contributes to roughly half of the total global costs [2]. People with dementia often exhibit behavioral and psychological symptoms of dementia (BPSD), including agitation, anxiety, irritability, depression, apathy, disinhibition, and changes in sleep or appetite [3]. More than 80% of people with dementia were affected by aggression, apathy, agitation, and depression, while around 30% experienced hallucinations and delusions [3,4]. Conceivably, BPSD adversely impacts the capacity to perform activities of daily living (ADL) and the maintenance of interpersonal relationships. This in turn results in significant stress and negative emotions for their family caregivers. For instance, caring for people with dementia has been associated with elevated levels of psychological distress and stress, as well as lower levels of self-efficacy, subjective well-being, and physical health [5]. Due to the significant level of psychological distress and stress involved in caregiving activities, approximately 50% of caregivers of people with dementia are susceptible to major depressive disorder [6]. A high level of caregiver stress is a risk factor for unhealthy behaviors, such as a sedentary lifestyle, inadequate dietary habits, and substance abuse, leading to an increased vulnerability to chronic illnesses among caregivers [7,8]. Caregiver burden is also linked to an increased risk of engaging in abusive behavior towards the care recipient [9]. It is therefore important to develop psychosocial interventions to promote the psychological well-being of both caregivers and people with dementia, as well as to alleviate the BPSD experienced by people with dementia.

Horticultural therapy (HT) has been proposed to be an effective intervention for improving the psychosocial well-being of people with dementia and their caregivers. HT is grounded in the Biophilia Hypothesis, which emphasizes the healing effects of human-plant interactions on physical, mental, and emotional domains [10]. According to the Attention Restoration Theory (ART), engaging with natural environments can restore cognitive fatigue by automatically capturing attention, thereby allowing the replenishment of directed attention capacity [11]. The Stress Recovery Theory (SRT) also proposes that interacting with nature can trigger physiological and psychological stress reduction and elicit positive affect, which in turn facilitates the replenishment of depleted psychological resources after stressful events [12,13]. For instance, HT establishes a sense of control, empowerment, and cooperation, which in turn can facilitate emotional stability when under a state of cognitive overload [14]. The physiological effects of HT on stress reduction in older adults are characterized by lower cortisol, heart rate, and blood pressure [15,16]. It has also been found that increased high α and low α wave activities after a single treatment [17]. Increasing α and theta activity in the brain's occipital lobes helped decrease anxiety and improve functioning in those with generalized anxiety disorder [18]. In addition, the social interactive nature of HT can facilitate social support for vulnerable groups [19].

More recently, HT has been extended in its application to people with dementia and demonstrated positive effects on BPSD, as well as enhancing physical and cognitive abilities, including memory and orientation, among people with dementia [20]. In the context of caregiving for people with dementia, HT can foster meaningful contact and social interaction between people with dementia and their caregivers [21,22]. A randomized controlled trial (RCT) demonstrated that HT can effectively reduce caregiving burden and promote the quality of life among caregivers of people with dementia [23]. Another RCT reported evidence of the protective effects of HT on the immune system and cognitive functioning, as well as its positive effects on reducing anxiety and promoting social interactions among older adults [24]. Several systematic reviews and meta-analyses have shown reliable effects of HT in improving cognitive functioning in people with dementia [25,26]. For instance, an improved capacity for autobiographical memory retrieval is often observed through engaging in horticultural activities [27]. In addition, HT also demonstrated benefits in alleviating BPSD in people with dementia, as suggested in multiple systematic reviews and meta-analyses [25,27-29].

Traditionally, HT is led by a registered therapist to achieve tailored rehabilitative or vocational goals. HT commonly includes participatory tasks such as cultivating, pruning, weeding, and planting flowers. Ornamental HT may also involve garden tours and nature viewing [16,22]. However, in densely populated areas, HT may not have the luxury of spaces to engage in home-based horticultural activities. The extreme climate in certain areas may create additional difficulties in engaging in outdoor horticultural activities. In addition, for caregivers who may be occupied with caregiving tasks, horticultural activities that require a significant amount of skills and time, such as monitoring the condition of various plants, may pose additional challenges for them [30].

While HT is promising in managing BPSD in people with dementia and improving the quality of life in both people with dementia and their caregivers, it is important to consider the environmental and spatial conditions in which HT is implemented, as these factors may inadvertently impact its

effectiveness. In this study, we proposed a novel approach to take advantage of technology to overcome the above limitations and to make horticultural activities more accessible for people with dementia and their caregivers, especially in densely populated areas. In recent years, there has been a growing body of research on the application of technology-assisted HT, such as using virtual reality to deliver HT (eg, VR Garden) [31]. However, there are certain limitations. One of the main drawbacks is the lack of real-world sensory experiences, such as smells and tactile sensations, which are inherent to traditional HT. Additionally, the use of virtual reality may lead to cybersickness, a condition characterized by symptoms similar to motion sickness, which can negatively impact the user experience and therapeutic outcomes [32]. To address these limitations, a hydroponic indoor grower can create a controllable, optimized environment for engaging in horticultural activities in a home-based HT for people with dementia and their caregivers. The hydroponic indoor grower also allows mobile app connectivity to track the progress of plant growth, which can guide people with dementia and their caregivers to engage in horticultural activities. Therefore, this pilot study aimed to address the following research objectives:

1. To examine the feasibility and acceptability of conducting a technology-enhanced home-based HT among people with dementia and their caregivers.
2. To investigate the preliminary effects of the intervention in (1) improving cognitive function and BPSD of people with dementia, (2) promoting positive caregiving experiences in family caregivers, and (3) reducing the caregivers' level of stress and depressive symptoms.

Methods

Study Design

To examine the feasibility and preliminary effects of the technology-enhanced home-based HT on improving the psychological well-being of people with dementia and their caregivers, we used an explanatory sequential multimethods approach. A single-group pre-post design is used for our quantitative assessment. The entire intervention lasted for 11 weeks, with outcome measures taken at baseline (T0) and post intervention (T1). A qualitative approach was used to explore the participants' experiences in-depth in the HT. Semistructured focus group interviews were conducted with participants after the intervention, with the aim of identifying the strengths, limitations, and difficulties experienced during the program.

Participants

Participants were recruited from a convenience sample of people with dementia and their caregivers in collaboration with 4 community centers operated by the nongovernmental organization (NGO) that provide gerontology services. Promotional posters were used to advertise the program in NGOs. The program allowed interested participants to enroll independently, while NGO staff reached out to interested individuals to provide further information and assistance. The Hong Kong version of the Montreal Cognitive Assessment-5-Minute (MoCA-5-min) [33] was administered by a research assistant in screening for stages of dementia in potential participants. Eligible participants were then enrolled based on the following inclusion and exclusion criteria. The inclusion criteria for people with dementia were as follows: (1) 65 years or above, (2) diagnosed with any type of dementia at the early to moderate stage, and (3) community-dwelling and living with family caregivers (ie, noninstitutionalized). The criteria for caregivers included the following: (1) at least 18 years old; (2) blood or by-marriage relatives (eg, spouses, siblings, children, and grandchildren) of a person who has been clinically diagnosed with dementia, regardless of its type; and (3) take up caring responsibilities ranging from physical aids (eg, transportation, financial assistance, personal hygiene, and decision-making) to emotional support, and provided most of the daily care and support for the people with dementia (daily contact of at least 4 hours). Participants were deemed ineligible if they had a current medical diagnosis with any acute mental illnesses that would otherwise prevent them from engaging in HT.

An a priori power analysis was performed using G*Power 3.1 [34], based on effect sizes in literature related to the effects of HT on psychological outcomes. A meta-analysis reported an effect size of 0.55 [35], the power analysis was formulated using a 1-tailed *t* test to detect a medium effect size with a power of 0.80 at a *P* value of .05, which resulted in a target sample size of 23. Considering an attrition rate of approximately 15% [36], it was necessary to include more than 27 dyads in the study.

Ethics Approval

Ethical approval for the study was obtained from the Institutional Review Board of The Hong Kong Polytechnic University (approval number HSEARS20220801002), and the trial was registered in ClinicalTrials.gov (NCT05577975). Prior to participation, all participants (both people with dementia and their caregivers) were required to provide informed consent. The process of obtaining informed consent from people with dementia adhered to the guidelines outlined by the Alzheimer's Society. A demonstration session was set up to explain the study's purpose, procedures, risks, benefits, and voluntary nature clearly to both people with dementia and their caregivers. During this session, the research assistant explained the consent form, demonstrated the equipment, and provided an overview of the participation process. People with dementia and their caregivers were allowed to try out some parts of the HT before signing the consent form. Only those people with dementia who were still willing to participate after this experience were allowed to sign the consent form. Ongoing assent was also obtained from participants with dementia throughout the study, and their willingness to participate was regularly monitored and respected.

Intervention

In the technology-enhanced home-based HT, the activities were also designed based on the Biophilia Hypothesis, with an emphasis on human-plant interactions. We further incorporated interactive components between people with dementia and their caregivers, in which people with dementia participated in the planting process alongside their caregivers. The caregivers assumed responsibility for decision-making and progress monitoring for the planting process, while people with dementia followed directions from caregivers and provided assistance

throughout the process. Examples of horticultural tasks performed by caregivers included formulating planting plans, adjusting optimal light conditions, strategizing fruit harvesting schedules, and devising culinary techniques for the harvest. Meanwhile, people with dementia executed tasks such as observation, flower watering, active participation in planting light regulation, and assisting caregivers with harvesting. According to the SRT and ART, these activities are posited to have cognitive and stress restoration effects. Based on the Sensory Integration Theory, these activities can mobilize the 5 senses through multisensory stimulation, tactile input, and olfactory and gustatory experiences [37,38], which can further enhance the cognitive functioning of people with dementia. Furthermore, the interactive components may enhance the dyadic relationship between people with dementia and their caregivers.

The technology-enhanced home-based HT was facilitated by a hydroponic indoor smart grower, which is a smart device that provides a greenhouse environment with adjustable LED grow lights and smart sensors (Figure 1). Features of the device include automatic watering and a removable water reservoir, as well as built-in smart sensors to detect nutrients, air temperature, humidity, light intensity, water temperature, and water level. The mobile app connectivity via Wi-Fi allows participants to monitor the growth conditions with notifications when water and nutrients need to be replenished, minimizing the probability of plant growing failure. The app can also customize planting plans for each type of plant and automatically adjust to suitable light intensity, resulting in faster plant growth and harvest. Additionally, the app can track the progress of plant growth and guide users to conduct horticultural activities.

Figure 1. Hydroponic indoor smart grower.

Upon the project's commencement, the hydroponic indoor smart grower was delivered to the enrolled participants. With reference to the optimal therapeutic dose between 100 and 500 minutes of HT for stress reduction [39], the intervention consisted of 3 weekly face-to-face (F-T-F) training sessions, each lasting 90 minutes. This was followed by 8 weeks of home-based horticultural activities (Table 1), with a total weekly time commitment of 1 hour. The program activities are presented in Table 2. A biweekly telephone call schedule was established to follow up on progress and program adherence. Additionally, a social media platform group was created for the participants to ask questions and share their horticultural experiences. The social media platform provided additional insight into participants' adherence to home-based horticultural activities.

Table 1. Intervention schedule of the technology-enhanced home-based horticultural therapy.

	F-T-F[a] sessions			Home-based horticultural activities with biweekly telephone follow-up							
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11
Duration (minutes)	90	90	90	60	60	60	60	60	60	60	60

[a]F-T-F: face-to-face.

Table 2. Program content of the technology-enhanced home-based horticultural therapy.

Week	Theme	Objectives	Reality orientation contents
4	Vision; Hearing; Smell	Stimulating the senses of vision, hearing, and smell; focusing on the appearance of the smart grower; encouraging people with dementia-caregiver interaction.	Focusing on LED[a] light differences according to certain modes of the smart grower; focusing on the water sound when the smart grower performs auto-watering system; smelling the seeds to see if different seeds have different smells; and sharing their sensory information to each other.
5	Smell; Touch; Motor	Stimulating the senses of smell and touch; enhancing visual-spatial skills and executive; promoting motor skills.	Smelling the leaves at the beginning stage of the plants; touching the plants (To see the texture, eg, hard/soft/smooth/rough/solid/fluid); trimming plants; taking photos to record plant growth.
6	Vision; Touch	Stimulating the senses of vision and touch; enhancing visual-spatial skills and executive functioning; enhancing cognitive ability through comparison.	Examining the pattern, shape, size, or height of the vegetables (eg, bok choy vs tomato) in day and night; touching different types of vegetables to see if the textures are different; sharing their sensory information to each other; taking photos to record plant growth.
7	Smell; Vision; Motor; Memory	Stimulating the senses of smell and touch; promoting motor functioning; enhancing visual-spatial skills and executive; enhancing cognitive functioning (memory) through comparison; encouraging people with dementia-caregiver interaction.	Comparing the shape, size, and height of different vegetables to see the differences in the growing process; smelling the leaves or stem of vegetables; pruning the vegetables; sharing and discussing what are the differences from the beginning; taking photos to record plant growth.
8	Motor; Touch; Decision making	Stimulating the senses of touch; promoting motor functioning; stimulating cognitive functioning through decision-making process; enhancing visual-spatial skills and executive.	Touching the leaves or stem of the plants or vegetables to see if the textures are different from the beginning to mid-late stage of planting; planning and discussing which plants to cut; taking photos to record plant growth.
9	Vision; Motor; Decision making	Stimulating the senses of vision; promoting motor functioning; stimulating cognitive functioning through decision-making process; enhancing visual-spatial skills and executive.	Counting the number of vegetables (eg, how many bak choy/tomato?); deciding which vegetables to cut; cutting suitable vegetables; taking photos to record plant growth.
10	Smell; Vision; Motor; Memory	Stimulating the senses of smell and touch; promoting motor functioning; enhancing visual-spatial skills and executive; enhancing cognitive functioning (memory) through comparison; encouraging people with dementia-caregiver interaction.	Comparing the shape, size, and height of different vegetables to see the differences in the growing process; smelling the leaves or stem of vegetables; pruning the vegetables; sharing and discussing what are the differences from the beginning; taking photos to record plant growth.
11	Motor; Vision; Decision-making	Stimulating the senses of vision; promoting motor functioning; stimulating cognitive functioning through decision-making process.	Deciding what type of dishes to cook (salads or pasta); cooking the decided dishes based on the vegetable; recording the odor and color of the dishes; taking photos to record plant growth.

[a]LED: light-emitting diode.

Measures

Overview

Measurements were taken at T0 and T1. The feasibility of the intervention was evaluated by means of attrition rate, attendance rate, and completion rate. The outcome measures for people with dementia included self-care abilities, cognitive function, behavioral symptoms, and happiness levels. Caregivers' outcomes consisted of positive caregiving experiences, perceived stress levels, depressive symptomatology, caregiver distress, and happiness levels.

People With Dementia Outcomes

The Barthel Index was used to measure the capacity to perform ADLs. Barthel Index comprised 10 ADLs, including feeding, bathing, grooming, dressing, using the toilet, transferring (moving from the bed to the chair and back), mobility (on level surfaces), climbing stairs, and controlling bowel and bladder functions [40]. Items are rated based on whether individuals can perform activities independently, with some assistance, or

are dependent (scored as 10, 5, or 0). The index yields a total score out of 100, with higher scores indicating greater levels of functional independence [40]. We categorized the Barthel Index scores as follows: 0-20 (total dependence), 21-60 (severe dependence), 61-90 (moderate dependence), 91-99 (mild dependence), and 100 (total independence) [41]. The Barthel Index had excellent test-retest reliability (intraclass correlation coefficient [ICC]≥0.962) [42].

The MoCA-5-min was used to measure the cognitive functions of people with dementia over the telephone, administered by social workers [33]. It comprises 4 domains: attention (immediate recall of 5 words), executive function/language (1-minute verbal fluency), orientation (6-item date and geographic orientation), and memory (delayed recall and recognition of 5 words learned in item 1). The total scores of the MoCA-5-min range from 0 to 30, with higher scores representing a higher level of cognitive function. The MoCA-5-min had an excellent test-retest reliability (ICC=0.89) [33].

BPSD of people with dementia was measured using the Neuropsychiatric Inventory Questionnaire (NPI-Q) [43]. The NPI-Q consisted of 12 neuropsychiatric symptoms, including delusions, hallucinations, agitation/aggression, irritability/depression, anxiety, euphoria/exuberance, apathy/indifference, disinhibition, irritability/instability, abnormal motor behavior, nocturnal behavioral disorders, and appetite/eating disorders. Each symptom is evaluated based on frequency (1=rarely to 4=very often, once or more per day) and severity (1=mild [causing little distress in the patient] to 3=severe [very disturbing to the patient and challenging to redirect]). A total score for each domain is calculated by multiplying the frequency and severity ratings. The overall NPI score is derived by summing the scores across all domains, with a higher score indicating more severe neuropsychiatric symptoms. The NPI-Q has demonstrated satisfactory psychometric properties [43].

The Visual Analogue Scale of Happiness was used to measure the happiness level of people with dementia [44]. The Visual Analogue Scale of Happiness is an analog scale containing 6 faces expressing emotions ranging from great happiness to deep sadness, and people with dementia need to identify which face best represents their mood. Subjects that chose the faces “Little Unhappy,” “Unhappy,” or “Very Unhappy” were prone to exhibit depressive symptoms [45]. The reliability and validity of the scale have been validated [46].

Caregivers Outcomes

The Positive Aspects of Caregiving Scale was used to measure caregivers’ positive caregiving experience [47]. It consists of 11 items with 2 subscales: enriching life and affirming self. Participants answered on a 5-point scale ranging from 0 (strongly disagree) to 4 (strongly agree). Scores range from 0 to 44, with higher scores indicating more positive self-perceptions of caregiving. The Positive Aspects of Caregiving Scale demonstrates high levels of internal consistency, with a Cronbach α of 0.89 among the family caregivers of people with dementia in Hong Kong [47].

The Chinese version of the Perceived Stress Scale was used to measure the stress level in caregivers [48]. It is a 10-item measurement rated on a 5-point Likert-type scale, ranging from 0 (never) to 4 (very often). The total score ranges from 0 to 40, with higher scores indicating higher levels of perceived stress. The Perceived Stress Scale showed acceptable levels of psychometric properties, which included an internal consistency reliability of Cronbach α of 0.85 and a test-retest reliability coefficient of 0.85 [49].

The Center for Epidemiological Studies Depression Scale was used to measure depressive symptoms in caregivers over a 1-week recall period [50]. It consists of 20 items and is rated on a 4-point Likert scale, ranging from 0 (rarely or none of the time) to 3 (most or almost all the time). Scores range from 0 to 60, with higher scores indicating greater depressive symptoms. The Center for Epidemiological Studies Depression Scale has demonstrated acceptable psychometric properties, including a test-retest reliability of 0.91 and an internal consistency of 0.86 [51].

The Neuropsychiatric Inventory-Distress Scale (NPI-D) was used to measure caregiving distress [52]. It assesses the psychological distress levels caused by neuropsychiatric symptoms, as reported in the NPI-Q, to the caregivers [43]. Participants rated their distress levels according to the severity of symptoms in people with dementia using a 6-point Likert scale, which ranges from 0 (not at all distressing) to 5 (very severely or extremely distressing). The Cronbach α coefficient for the Chinese version of the NPI-D was 0.72 [53].

Perceived happiness was measured using the Subjective Happiness Scale (SHS) [54]. The SHS consists of 4 items; 2 of these assess the current state of happiness, including absolute happiness ratings and relative happiness ratings compared with peers. The remaining 2 items measure trait happiness by asking participants to indicate the extent to which they identify with descriptions of both happy and unhappy individuals. The SHS is rated on a 7-point Likert-type scale, ranging from negative to positive. Responses to the 4 items are averaged, yielding a range from 1 to 7. A higher score indicates greater levels of happiness. The Cronbach α for the Chinese version of the SHS was 0.82, and the test-retest reliability was 0.70 [55].

Feasibility Outcomes

Intervention feasibility was assessed using the completion rate, attendance rate, and attrition rate. The attrition rate was calculated by dividing the number of dropouts by the total number of enrolled subjects. The attendance rate was determined by averaging the individual attendance rates, which were calculated by dividing the number of F-T-F sessions attended by the total number of F-T-F sessions. The completion rate was defined as the percentage of participants who attended more than 80% of the sessions, which included home-based horticultural activities and totaled more than 600 minutes.

Quantitative Analysis

Data analysis was performed using IBM SPSS (version 23). We first explored patterns of sociodemographic characteristics of the participants. Then, paired-sample *t* tests were used to compare the outcomes before and after the intervention. Unless

otherwise specified, the statistical analyses performed did not exceed a 2-tailed α level of 0.05.

Qualitative Analysis

Given that caregivers of people with dementia experience varying levels of distress and may respond differently to the intervention, we adopted stratified purposive sampling based on this criterion to enhance the representativeness of the data. The caregiver's level of distress was assessed using NPI-D, with mild distress categorized as 0-13 points (77.78%), moderate distress as 14-26 points (16.67%), and severe distress as >27 points (5.56%) [52]. Participants for the qualitative study were selected through stratified sampling based on these different levels of distress. A total of 4 semistructured focus groups were conducted with 18 participants. A trained research assistant led the focus groups, asking participants to share their experiences using the smart growers, as well as their interactions with people with dementia during the program. A probing question—"What is your experience related to the use of the device in interacting with your relative?"—was used to facilitate the discussions. The focus group interviews were recorded and transcribed with the participants' consent.

Thematic analysis was used to analyze the transcribed verbatim; the authors read the texts, generated initial codes, and developed potential themes [56]. The process concluded with data saturation, when no new findings were discovered. Emergent themes were discussed and agreed upon by the researchers. The comprehensiveness of the content analysis was ensured through digital recordings and verbatim transcriptions, which were independently completed and then discussed by 2 researchers. The emergent themes, validated through researcher consensus after thorough discussion, exemplify the methodological rigor underlying the study's credibility.

Results

Quantitative Findings

A total of 50 dyads of people with dementia and their family caregivers were invited to participate in the study, and 38 dyads met the selection criteria and joined the study. The attrition rate was 2.63%, with 1 participant dropping out due to behavioral issues after the first F-T-F training session. The attendance rate for F-T-F sessions was 94.64%. A total of 36 out of 38 dyads attended the first session, while 35 out of 37 dyads attended both the second and third sessions. The completion rate was 83.78%, as 31 out of 37 participants achieved a participation rate of 80% or higher (>600 minutes).

Sample characteristics of people with dementia are presented in Table 3. The majority of participants were male (21/38, 55.3%), with a mean age of 81.8 (SD 8.4) years. Regarding the self-care abilities of the participants, only 21.1% (n=8) exhibited no dependence, while the majority displayed mild to moderate dependence.

Table 3. Sociodemographic characteristics of people with dementia (N=38).

Variables	Values
Age (in years), mean (SD)	81.8 (8.4)
Age (in years), range	60-102
Sex , n (%)	
Male	21 (55.3)
Female	17 (44.7)
Marital status , n (%)	
Married	23 (60.5)
Divorced/Widowed	15 (39.5)
Self-care abilities , n (%)	
Total independence	8 (21.1)
Mild dependence	25 (65.8)
Moderate dependence	5 (13.2)
Severe dependence	0 (0.0)
Total dependence	0 (0.0)

Demographic characteristics of caregivers are summarized in Table 4. Female caregivers accounted for 73.7%, with a mean age of 58.4 (SD 8.8) years. The average duration of care provided to people with dementia was 4.7 (SD 0.3) hours per day.

Table 4. Sociodemographic characteristics of caregivers (N=38).

Variables	Values
Age (in years), mean (SD)	58.4 (8.8)
Age (in years), range	34-76
Sex, n (%)	
Male	10 (26.3)
Female	28 (73.7)
Marital status, n (%)	
Single	2 (5.3)
Married	30 (78.9)
Divorced/widowed	6 (15.8)
Educational level, n (%)	
Primary	10 (26.3)
Secondary	20 (52.6)
College or above	8 (21.1)
Employment status, n (%)	
Unemployed/retired	20 (52.6)
Employed	18 (47.4)
Monthly household income, n (%)	
Less than 6000 HKD[a]	11 (28.9)
6000-9999 HKD	8 (21.1)
10,000-14,999 HKD	10 (26.3)
15,000-19,999 HKD	7 (18.4)
20,000 HKD or above	2 (5.3)
Relationship with people with dementia, n (%)	
Spouse	20 (52.6)
Children/in-laws	18 (47.4)
Living arrangement with people with dementia, n (%)	
Same household	23 (60.5)
Different household	15 (39.5)
Duration of care provided to people with dementia, in hours/day (hours), mean (SD)	4.7 (0.3)

[a]HKD: Hong Kong Dollar (1 HKD=US $ 0.13).

For people with dementia, there were significant pre-post improvements in cognitive functions and behavioral and BPSD, although the results did not reach statistical significance. Furthermore, we observed a significant improvement in the happiness level of people with dementia from T0 (mean 7.67, SD 1.40) to T1 (mean 8.79, SD 1.02; P<.01; Table 5).

Table 5. Outcome measures of the technology-enhanced home-based horticultural therapy for people with dementia (N=37, one participant dropped out).

Variables	T0[a], mean (SD)	T1[b], mean (SD)	*P* value	95% CI
Happiness	7.67 (1.40)	8.79 (1.02)	<.01	–1.75 to –0.46
Cognitive functions	14.07 (6.58)	14.36 (6.55)	.61	–1.41 to 0.84
BPSD[c]	10.03 (8.26)	7.72 (5.22)	.20	–1.29 to 5.91

[a]T0: preintervention.

[b]T1: postintervention.

[c]BPSD: behavioral and psychological symptoms of dementia.

Regarding caregivers, at a descriptive level, we observed pre-post improvements in perceived stress, depressive symptoms, and happiness, although the differences did not reach a statistically significant level. Specifically, we observed significant improvements in caregiver distress from T0 (mean 12.93, SD 9.40) to T1 (mean 7.83, SD 8.01; $P<.05$; Table 6).

Table 6. Outcome measures of the technology-enhanced home-based horticultural therapy for caregivers (N=37, one participant dropped out).

Variables	T0[a], mean (SD)	T1[b], mean (SD)	*P* value	95% CI
Stress	25.72 (6.17)	25.66 (6.00)	.96	–2.89 to 3.03
Depressive symptoms	16.79 (10.70)	17.03 (9.73)	.92	–5.28 to 4.80
Positive aspects of caregiving	31.34 (8.23)	31.00 (10.29)	.89	–4.86 to 5.55
Caregiver distress	12.93 (9.40)[c]	7.83 (8.01)[c]	<.05[c]	0.16 to 10.05[c]
Happiness	6.74 (3.93)	7.08 (4.37)	.41	–1.42 to 3.35

[a]T0: preintervention.

[b]T1: postintervention.

[c]Significant effects.

Qualitative Findings

Overview

A total of 18 caregivers joined the focus group. The majority of them were either children (n=12) or spouses (n=6) of the care recipients. Overall, the qualitative findings from the focus interviews can be summarized under the themes of (1) perceived benefits and (2) perceived challenges of the technology-enhanced home-based HT.

Perceived Benefits

Under the theme of perceived benefits, we further identified four subthemes: (1) improved emotional well-being, (2) improved relationship with people with dementia, (3) increased social network among caregivers, and (4) reinforced autobiographical memory retrieval and storytelling abilities.

Improved Emotional Well-Being

Caregivers positively recalled how they had enjoyed engaging in horticultural activities at home using the smart grower. They shared their excitement at witnessing the growing process of the vegetables. One participant shared that growing vegetables alleviated her stress, as it had become a meaningful goal in her life. Additionally, she appreciated the colorful lighting of the machine, which they found to be a decorative element in their homes.

> *Seeing the growth of vegetables is like looking after a baby. The first thing I do every day is to take photos of the vegetables… I also feel happy when I see the attractive symbols on the machine, such as the stars and mountain symbols.* [P3]

The caregivers were pleased that they could enjoy the fruits of their labor. The vegetables they grew, such as carrots and lettuce, were used to create delicious dishes. They felt a sense of pride and satisfaction after receiving compliments from friends and family members who saw photos of their harvest.

Similar perspectives were also evident in people with dementia. They felt happiness and excitement about the growing process, especially during the harvest. As one caregiver shared,

> *My father was also involved in the growing of lettuce…He looked at the smart grower and watered the plant every day. He felt happy when he saw the harvest in the smart grower…He could feel the nature at home.* [P10]

Another caregiver noted that while her husband could not actively participate in the growing process, he enjoyed watching it, particularly the colorful lighting on the smart growing machine.

Improved Relationship With People With Dementia

Caregivers found that their relationship with people with dementia improved through home-based horticultural activities. As one caregiver explained,

> *I put the smart grower near the window, and my husband (people with dementia) will help to water the plant…We are taking care of the plant together…It is just like a toy for us.* [P4]

These activities offered a way for them to collaborate and grow the vegetables together.

Increase Social Network Among Caregivers

Caregivers also reported increased interaction with other caregivers through the home horticultural activities. They were able to ask questions about the smart grower and share photos of their harvest in WhatsApp groups. This facilitated the formation of new friendships, as they had a common topic to discuss.

I uploaded photos to the WhatsApp group and others posted pictures on the group. So, I see how other people's plants are growing and talk to people.(P1) We have set a WhatsApp groups. When I don't understand sometimes, I will ask in the group. The student assistants will help me in the group and other caregivers will also answer and we will discuss together in the group. [P9]

The increased social interactions were corroborated by our WhatsApp group records. Participants actively shared recipes for their harvests with each other. Some caregivers displayed their stir-fried Chinese cabbage, while others shared recipes for garlic Chinese cabbage and methods for cooking baby turnip

with chieh-qua. More noteworthy was how caregivers enthusiastically shared their planting progress and complimented each other's growing vegetables. These positive interactions led to the formation of new friendships among participants.

Reinforced Autobiographical Memory Retrieval and Storytelling Abilities

One caregiver noted that the smart grower offered an opportunity for people with dementia to reminisce about the past. The people with dementia had experience growing crops in their homeland during childhood, and the home horticultural activities revived those childhood memories. This opportunity facilitated the sharing of historical anecdotes and the exercise of cognitive abilities, such as autobiographical memory retrieval.

> *Dad is interested in gardening, as growing vegetables with this machine reminded him of his country gardening days, and every now and then he would tell us about his previous experiences.* [P13]

Difficulties

While the horticultural activities generally had positive impacts on people with dementia and their caregivers, there were also some perceived challenges. These challenges can be summarized into two subthemes: (1) operational difficulties in using the smart grower, and (2) light disturbance from the smart grower.

Operational Difficulties in Using the Smart Grower

Some caregivers found the smart grower not easy to operate due to its sophisticated design featuring many high-tech components. Issues such as Wi-Fi connection problems were encountered during use. For people with dementia with mild to moderate dependence, the complexity of the operations seemed to diminish their patience and motivation, presenting a barrier to their active participation in the entire growth process.

> *Sometimes, the elder have no caregivers to help them. So, they can't do it by themselves because they don't understand English very well. For example, the word 'reset' they can't understand. Also, they would not know what to do if the Wi-Fi disconnected.* [P13]

Light Disturbance from the Smart Grower. Additionally, other caregivers expressed concerns about the lighting emitted by the smart grower. They reported using covers or towels to minimize light disturbance at night, indicating a need for better light management features in the device.

> *The machine emits light and is very disruptive to sleep. Sleeping in a very small house is really affected. When sleeping in the living room, the light turns red and is a little scary. It may have some effect on plants, but not sure. The next day it gives off normal light again.* [P14]

Discussion

Principal Findings

This pilot study aimed to investigate the feasibility of technology-enhanced home-based HT in improving the psychological well-being of people with dementia and their caregivers. Overall, our quantitative results suggested that technology-enhanced home-based HT is a feasible dyadic intervention for both people with dementia and their caregivers. The overall attrition rate was low (2.63%), with a satisfactory completion rate (83.78%). These findings indicate that caregivers and people with dementia are receptive to using technology for engaging in horticultural activities. Further qualitative findings revealed that both people with dementia and their caregivers were excited about using the smart growing machine, as it provided a novel experience for them. They also welcomed the interactive horticultural activities and found that their relationship with each other improved as a result.

The use of technology-enhanced HT has progressively become an increasingly popular approach to delivering HT. For example, virtual reality–based HT has recently been demonstrated as a viable and effective approach in promoting positive psychological outcomes in older adults, such as enhancing self-esteem and mastery [57]. Technology-enhanced HT can mitigate limitations posed by factors such as venue and weather conditions, offering a more flexible and convenient therapeutic approach. In contrast to a virtual reality–based HT, our technology-enhanced home-based HT not only exhibits the benefits of using technologies but also allows participants to touch and feel plants in reality, which can be more relatable to the feeling of nurturing life. According to the Biophilic Hypothesis, actual human-plant interactions play a vital role in exerting the therapeutic effects of HT [58].

Our quantitative results suggested a significant reduction in caregiver distress, aligning with previous studies that used a traditional delivery mode of HT. Previous research has demonstrated that horticultural activities can alleviate the caregiving burden [23]. Our qualitative findings further suggested that home-based horticultural activities may simultaneously enable people with dementia to engage in meaningful and socially interactive experiences with their caregivers, potentially improving their relationships and alleviating caregiver distress. Moreover, our findings indicated a significant increase in happiness levels among people with dementia. These findings also corroborated previous studies using traditional HT that reported a positive emotional impact on people with dementia [59]. Our qualitative findings provided further insights that the heightened levels of subjective well-being could be attributed to participants' enthusiasm for observing plant growth and acquiring a sense of satisfaction through harvesting activities.

Overall, our findings are in line with the ART, in that perceiving plants in their natural environment has the potential to enhance attention restoration and increase persuasive effects [11,60]. Our findings are also in line with SRT, suggesting that HT is associated with positive affect, which in turn facilitates stress recovery [12,61]. Although we did not observe a significant decrease in perceived stress levels, we found an overall improvement in emotional well-being among caregivers. For example, an increase in emotional well-being is inversely correlated with elevated levels of stress and anger [62]. Consistently, HTs can reduce cortisol levels through engagement in horticultural activities, such as self-harvesting processes, and the consumption of harvested fruits and vegetables [15]. The use of social media groups also facilitated social interactions

among caregivers and provided a source of support. The mutual encouragement and compliments about each other's vegetables, along with sharing recipes for their harvest created positive interactions that extended beyond the initial horticultural activities, offering an additional benefit of alleviating caregiving stress [63-65].

Our qualitative findings suggested that some participants reminisced about their past experiences in farming during the activities. Reminiscing about past events requires the recall of autobiographical memory [66,67], and actively externalizing autobiographical memory for later retrieval (ie, reminiscing about past events) can enhance memory performance and afford greater control over memory encoding processes [68]. While previous studies have suggested that engagement in horticultural activities can enhance the cognitive abilities of people with dementia [21,22], we did not detect a significant pre-post improvement in the cognitive functions of people with dementia. It is plausible that because the MoCA is a global measure of cognition, it may partially explain why we did not detect a significant pre-post improvement in MoCA scores, despite an increasing trend descriptively, as people with dementia primarily improved in memory abilities. However, this interpretation should be made cautiously due to our small sample size. Future studies should incorporate a larger-scale RCT to further investigate the cognitive benefits of the technology-enhanced home-based HT.

Limitations

This study had several limitations. Due to the pilot nature of the research to establish the feasibility of a technology-enhanced home-based HT, a control group was not employed. Therefore, a full-scale RCT is encouraged to provide further confirmatory evidence about the effectiveness of this intervention. Furthermore, there were only 3 F-T-F sessions provided to offer training and support to the participants. This dosage may have been insufficient. In the qualitative findings, caregivers reported having technical difficulties in using the smart grower at home, such as Wi-Fi connection issues. Previous studies have found similar observations that the older population may have a lower level of readiness for home technology [69]. Therefore, the complexity of the device use may be a significant barrier to the engagement of the older population, particularly for people with dementia [70]. Consequently, it is conceivable that the program outcomes could be further enhanced with more F-T-F sessions and better integrated technical support for people with dementia and their caregivers. A more comprehensive training and support system may help address the operational difficulties encountered with the smart grower technology.

Conclusion

The findings of this pilot study support the feasibility and offer preliminary evidence of using a hydroponic indoor grower to conduct technology-enhanced home-based HT to improve the psychological well-being of people with dementia and their caregivers. There were positive effects on the emotional well-being of people with dementia and a decrease in caregiver distress. The qualitative findings further suggested improved relationships between people with dementia and caregivers. Additionally, there were potential benefits of this intervention in improving the autobiographical memory retrieval capacity of people with dementia, indicating the cognitive benefits of this intervention. Future studies are encouraged to incorporate a larger and more diverse sample to further evaluate the sustainability of the positive psychological benefits of the technology-enhanced home-based HT for people with dementia and their caregivers.

Acknowledgments

This study received funding from Bright Future Charitable Foundation, Hong Kong, China. This publication reflects only the authors' views, and the Bright Future Charitable Foundation is not liable for any use that may be made of the information contained therein.

Authors' Contributions

PK, JL, AW, DC, HL, and FW conceptualized the research design. AT and HZT conducted the interview and assisted in data collection. AT and HZT conducted the qualitative analyses. PK, AT, and HZT performed statistical analyses and interpreted the data. PK, PK, AT, and HZT drafted the manuscript. All authors read and approved the final manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Shin JH. Dementia epidemiology fact sheet 2022. Ann Rehabil Med 2022;46(2):53-59 [FREE Full text] [doi: 10.5535/arm.22027] [Medline: 35508924]
2. Wimo A, Seeher K, Cataldi R, Cyhlarova E, Dielemann JL, Frisell O, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023;19(7):2865-2873 [FREE Full text] [doi: 10.1002/alz.12901] [Medline: 36617519]
3. Scales K, Zimmerman S, Miller SJ. Evidence-based nonpharmacological practices to address behavioral and psychological symptoms of dementia. Gerontologist 2018;58(suppl_1):S88-S102 [FREE Full text] [doi: 10.1093/geront/gnx167] [Medline: 29361069]

4. Holroyd S, Currie L, Wooten GF. Prospective study of hallucinations and delusions in parkinson's disease. J Neurol Neurosurg Psychiatry 2001 Jun;70(6):734-738 [FREE Full text] [doi: 10.1136/jnnp.70.6.734] [Medline: 11385005]
5. Ballard C, Corbett A, Chitramohan R, Aarsland D. Management of agitation and aggression associated with alzheimer's disease: controversies and possible solutions. Curr Opin Psychiatry 2009;22(6):532-540. [doi: 10.1097/YCO.0b013e32833111f9] [Medline: 19696673]
6. Huang SS. Depression among caregivers of patients with dementia: associative factors and management approaches. World J Psychiatry 2022;12(1):59-76 [FREE Full text] [doi: 10.5498/wjp.v12.i1.59] [Medline: 35111579]
7. Hoffman D, Zucker H. A call to preventive action by health care providers and policy makers to support caregivers. Prev Chronic Dis 2016;13:E96 [FREE Full text] [doi: 10.5888/pcd13.160233] [Medline: 27442996]
8. Culberson JW, Kopel J, Sehar U, Reddy PH. Urgent needs of caregiving in ageing populations with alzheimer's disease and other chronic conditions: support our loved ones. Ageing Res Rev 2023;90:102001. [doi: 10.1016/j.arr.2023.102001] [Medline: 37414157]
9. Wang M, Sun H, Zhang J, Ruan J. Prevalence and associated factors of elder abuse in family caregivers of older people with dementia in central China cross-sectional study. Int J Geriatr Psychiatry 2019;34(2):299-307. [doi: 10.1002/gps.5020] [Medline: 30370657]
10. Niklasson J. Horticultural therapy for homeless people. SLU University Library. 2007. URL: https://stud.epsilon.slu.se/12506/ [accessed 2017-10-26]
11. Kaplan S. The restorative benefits of nature: toward an integrative framework. J Environ Psychol 1995;15(3):169-182. [doi: 10.1016/0272-4944(95)90001-2]
12. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol 1991;11(3):201-230. [doi: 10.1016/s0272-4944(05)80184-7]
13. Vacher P, Filaire E, Mourot L, Nicolas M. Stress and recovery in sports: effects on heart rate variability, cortisol, and subjective experience. Int J Psychophysiol 2019;143:25-35. [doi: 10.1016/j.ijpsycho.2019.06.011] [Medline: 31255740]
14. Kenmochi T, Kenmochi A, Hoshiyama M. Effects of horticultural therapy on symptoms and future perspective of patients with schizophrenia in the chronic stage. Journal of Therapeutic Horticulture 2019;29(1):1-10 [FREE Full text]
15. Yun J, Yao W, Meng T, Mu Z. Effects of horticultural therapy on health in the elderly: a review and meta-analysis. Z Gesundh Wiss 2023;32(10):1-27 [FREE Full text] [doi: 10.1007/s10389-023-01938-w] [Medline: 37361284]
16. Nicholas SO, Giang AT, Yap PLK. The effectiveness of horticultural therapy on older adults: a systematic review. J Am Med Dir Assoc 2019;20(10):1351.e1-1351.e11. [doi: 10.1016/j.jamda.2019.06.021] [Medline: 31402135]
17. Hassan A, Qibing C, Tao J. Physiological and psychological effects of gardening activity in older adults. Geriatr Gerontol Int 2018;18(8):1147-1152. [doi: 10.1111/ggi.13327] [Medline: 29626378]
18. Dadashi M, Birashk B, Taremian F, Asgarnejad AA, Momtazi S. Basic Clin Neurosci 2015;6(1):14-20 [FREE Full text] [Medline: 27504152]
19. Sempik J, Rickhuss C, Beeston A. The effects of social and therapeutic horticulture on aspects of social behaviour. Br J Occup Ther 2014;77(6):313-319. [doi: 10.4276/030802214x14018723138110]
20. Lin Y, Lin R, Liu W, Wu W. Effectiveness of horticultural therapy on physical functioning and psychological health outcomes for older adults: a systematic review and meta-analysis. J Clin Nurs 2022;31(15-16):2087-2099. [doi: 10.1111/jocn.16095] [Medline: 34694042]
21. Jarrott SE, Kwack HR, Relf D. An observational assessment of a dementia-specific horticultural therapy program. HortTechnology 2002;12(3):403-410. [doi: 10.21273/horttech.12.3.403]
22. Ferrini F. Horticultural therapy and its effect on people's health. Adv Hortic Sci 2003;17(2):77-87 [FREE Full text]
23. Kim YH, Park CS, Bae H, Lim EJ, Kang KH, Lee ES, et al. Horticultural therapy programs enhancing quality of life and reducing depression and burden for caregivers of elderly with dementia. J People Plants Environ 2020;23(3):305-320. [doi: 10.11628/ksppe.2020.23.3.305]
24. Ng KST, Sia A, Ng MKW, Tan CTY, Chan HY, Tan CH, et al. Effects of horticultural therapy on asian older adults: a randomized controlled trial. Int J Environ Res Public Health 2018;15(8):1705 [FREE Full text] [doi: 10.3390/ijerph15081705] [Medline: 30096932]
25. Zhao Y, Liu Y, Wang Z. Effectiveness of horticultural therapy in people with dementia: a quantitative systematic review. J Clin Nurs 2022;31(13-14):1983-1997. [doi: 10.1111/jocn.15204] [Medline: 32017241]
26. Tu HM, Chiu PY. Meta-analysis of controlled trials testing horticultural therapy for the improvement of cognitive function. Sci Rep 2020;10(1):14637 [FREE Full text] [doi: 10.1038/s41598-020-71621-7] [Medline: 32884098]
27. Lu LC, Lan SH, Hsieh YP, Yen YY, Chen JC, Lan SJ. Horticultural therapy in patients with dementia: a systematic review and meta-analysis. Am J Alzheimers Dis Other Demen 2020;35:1533317519883498 [FREE Full text] [doi: 10.1177/1533317519883498] [Medline: 31690084]
28. Whear R, Coon JT, Bethel A, Abbott R, Stein K, Garside R. What is the impact of using outdoor spaces such as gardens on the physical and mental well-being of those with dementia? a systematic review of quantitative and qualitative evidence. J Am Med Dir Assoc 2014;15(10):697-705 [FREE Full text] [doi: 10.1016/j.jamda.2014.05.013] [Medline: 25037168]

29. Xu M, Lu S, Liu J, Xu F. Effectiveness of horticultural therapy in aged people with depression: a systematic review and meta-analysis. Front Public Health 2023;11:1142456 [FREE Full text] [doi: 10.3389/fpubh.2023.1142456] [Medline: 36969640]
30. Kamioka H, Tsutani K, Yamada M, Park H, Okuizumi H, Honda T, et al. Effectiveness of horticultural therapy: a systematic review of randomized controlled trials. Complement Ther Med 2014;22(5):930-943. [doi: 10.1016/j.ctim.2014.08.009] [Medline: 25440385]
31. Xu T, Liao D, Peng X, Liang G, Li Y, Zheng X. Design of virtual reality scenes on horticultural therapy treatment. 2012 Presented at: International Conference on Human-Computer Interaction; 2022 June 26; Heidelberg. [doi: 10.15695/vurj.v8i0.3644]
32. Rebenitsch L, Owen C. Review on cybersickness in applications and visual displays. Virtual Reality 2016;20(2):101-125. [doi: 10.1007/s10055-016-0285-9]
33. Wong A, Nyenhuis D, Black SE, Law LS, Lo ES, Kwan PW, et al. Montreal cognitive assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration. Stroke 2015;46(4):1059-1064. [doi: 10.1161/strokeaha.114.007253]
34. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009;41(4):1149-1160. [doi: 10.3758/brm.41.4.1149]
35. Pan iru I, Ronaldson A, Sima N, Dregan A, Sima R. The impact of gardening on well-being, mental health, and quality of life: an umbrella review and meta-analysis. Syst Rev 2024;13(1):45 [FREE Full text] [doi: 10.1186/s13643-024-02457-9] [Medline: 38287430]
36. Siu AMH, Kam M, Mok I. Horticultural therapy program for people with mental illness: a mixed-method evaluation. Int J Environ Res Public Health 2020;17(3) [FREE Full text] [doi: 10.3390/ijerph17030711] [Medline: 31979057]
37. Smith Roley S, Mailloux Z, Miller-Kuhaneck H, Glennon TJ. Understanding Ayres' sensory integration. Sacred Heart University 2007;12(17):CE1-CE8 [FREE Full text]
38. Yun HS, Yun SY, Choi BJ. Effects of horticultural activities designed to stimulate five senses on the sensory development of children. J People Plants Environ 2018;21(5):369-378. [doi: 10.11628/ksppe.2018.21.5.369]
39. Lu S, Liu J, Xu M, Xu F. Horticultural therapy for stress reduction: a systematic review and meta-analysis. Front Psychol 2023;14:1086121 [FREE Full text] [doi: 10.3389/fpsyg.2023.1086121] [Medline: 37564307]
40. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965;14:61-65. [Medline: 14258950]
41. Filipska-Blejder K, Ślusarz R. Characteristics of the barthel scale in terms of neurological care. PNN 2022;11(4):180-184. [doi: 10.15225/pnn.2022.11.4.6]
42. Dos Santos Barros V, Bassi-Dibai D, Guedes CLR, Morais DN, Coutinho SM, de Oliveira Simões G, et al. Barthel Index is a valid and reliable tool to measure the functional independence of cancer patients in palliative care. BMC Palliat Care 2022;21(1):124 [FREE Full text] [doi: 10.1186/s12904-022-01017-z] [Medline: 35820921]
43. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, et al. Validation of the NPI-Q, a brief clinical form of the neuropsychiatric inventory. J Neuropsychiatry Clin Neurosci 2000;12(2):233-239. [doi: 10.1176/jnp.12.2.233] [Medline: 11001602]
44. Matsubayashi K, Kimura S, Iwasaki T, Okumiya K, Hamada T, Fujisawa M, et al. [Evaluation of subjective happiness in the elderly using a visual analogue scale of happiness to analyze the effect of life style and neurobehavioral function on subjective happiness]. Nihon Ronen Igakkai Zasshi 1992;29(11):817-822 [FREE Full text] [doi: 10.3143/geriatrics.29.817] [Medline: 1491476]
45. César KG, Brucki SMD, Takada LT, Nascimento LFC, Gomes CMS, Almeida MCS, et al. Performance of the visual analogue scale of happiness and of the cornell scale for depression in dementia in the tremembé epidemiological study, Brazil. Dement Neuropsychol 2014;8(4):389-393 [FREE Full text] [doi: 10.1590/S1980-57642014DN84000014] [Medline: 29213931]
46. Matsubayashi K, Kimura S, Iwasaki T, Okumiya S, Hamada T, Fujisawa M, et al. [Evaluation of subjective happiness in the elderly using a visual analogue scale of happiness in correlation with depression scale]. Nihon Ronen Igakkai Zasshi 1992;29(11):811-816 [FREE Full text] [doi: 10.3143/geriatrics.29.811] [Medline: 1491475]
47. Lou VW, Lau BH, Cheung KS. Positive aspects of caregiving (PAC): scale validation among Chinese dementia caregivers (CG). Arch Gerontol Geriatr 2015;60(2):299-306. [doi: 10.1016/j.archger.2014.10.019] [Medline: 25488014]
48. Cohen S, Kamarck T, Mermelstein R. Perceived stress scale. APA PsycTests 1983:10. [doi: 10.1037/t02889-000]
49. Leung DY, Lam TH, Chan SS. Three versions of perceived stress scale: validation in a sample of Chinese cardiac patients who smoke. BMC Public Health 2010;10:513 [FREE Full text] [doi: 10.1186/1471-2458-10-513] [Medline: 20735860]
50. Chin WY, Choi EP, Chan KT, Wong CK. The psychometric properties of the center for epidemiologic studies depression scale in Chinese primary care patients: factor structure, construct validity, reliability, sensitivity and responsiveness. PLoS One 2015;10(8):e0135131 [FREE Full text] [doi: 10.1371/journal.pone.0135131] [Medline: 26252739]
51. Miller WC, Anton HA, Townson AF. Measurement properties of the CESD scale among individuals with spinal cord injury. Spinal Cord 2008;46(4):287-292. [doi: 10.1038/sj.sc.3102127] [Medline: 17909558]

52. Kaufer DI, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, et al. Assessing the impact of neuropsychiatric symptoms in alzheimer's disease: the neuropsychiatric inventory caregiver distress scale. J Am Geriatr Soc 1998;46(2):210-215. [doi: 10.1111/j.1532-5415.1998.tb02542.x] [Medline: 9475452]
53. Wang T, Xiao S, Li X, Wang H, Liu Y, Su N, et al. Reliability and validity of the Chinese version of the neuropsychiatric inventory in mainland China. Int J Geriatr Psychiatry 2012;27(5):539-544. [doi: 10.1002/gps.2757] [Medline: 21769938]
54. Lyubomirsky S, Lepper HS. A measure of subjective happiness: preliminary reliability and construct validation. Soc Indic Res 1999;46(2):137-155. [doi: 10.1023/A:1006824100041]
55. Nan H, Ni MY, Lee PH, Tam WW, Lam TH, Leung GM, et al. Psychometric evaluation of the Chinese version of the subjective happiness scale: evidence from the Hong Kong FAMILY cohort. Int J Behav Med 2014;21(4):646-652 [FREE Full text] [doi: 10.1007/s12529-014-9389-3] [Medline: 24515396]
56. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
57. Fan CC, Choy CS, Huang CM, Chih PS, Lee CC, Lin FH, et al. The effects of a combination of 3D virtual reality and hands-on horticultural activities on mastery, achievement motives, self-esteem, isolation and depression: a quasi-experimental study. BMC Geriatr 2022;22(1):744 [FREE Full text] [doi: 10.1186/s12877-022-03431-7] [Medline: 36096746]
58. Kaplan R. Some psycholog Ical benefits of gardening. Environ Behav 1973;5(2):145-162. [doi: 10.1177/001391657300500202]
59. Gigliotti CM, Jarrott SE, Yorgason J. Harvesting health: effects of three types of horticultural therapy activities for persons with dementia. Dementia 2004;3(2):161-180. [doi: 10.1177/1471301204042335]
60. Ulrich RS, Simons RF, Losito BD, Fiorito E, Miles MA, Zelson M. Stress recovery during exposure to natural and urban environments. J Environ Psychol 1991;11(3):201-230. [doi: 10.1016/s0272-4944(05)80184-7]
61. Zuckerman M. Development of a situation-specific trait-state test for the prediction and measurement of affective responses. J Consult Clin Psychol 1977;45(4):513-523. [doi: 10.1037//0022-006x.45.4.513] [Medline: 886036]
62. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. Lancet 2015;385(9968):640-648 [FREE Full text] [doi: 10.1016/S0140-6736(13)61489-0] [Medline: 25468152]
63. Hui Lin H, Mu Chuan L, Wan Lan C. Benefits of horticultural therapy on emotional and perceived stress of family caregivers: a pilot study. Research in Applied Psychology 2017;67:5-53.
64. Jueng RN, Lin CY, Huang YH. Systematic review on the positive mental health impact of older adults participation in horticultural activities in long term care facilities. Horticulturae 2023;9(10):1076. [doi: 10.3390/horticulturae9101076]
65. Houben M, Van Engen V, Kenning G, Brankaert R. Smile: capturing and sharing personal photos to stimulate social relations and support self-identity in dementia. New York: Springer; 2021.
66. Lu X, Kelly MO, Risko EF. Offloading information to an external store increases false recall. Cognition 2020;205:104428. [doi: 10.1016/j.cognition.2020.104428] [Medline: 32863020]
67. Janssen SMJ. Introduction to the cognitive abilities account for the reminiscence bump in the temporal distribution of autobiographical memory. Psychol Rep 2020;123(1):12-42. [doi: 10.1177/0033294119843221] [Medline: 31039675]
68. Kelly MO, Risko EF. Offloading memory: serial position effects. Psychon Bull Rev 2019;26(4):1347-1353. [doi: 10.3758/s13423-019-01615-8] [Medline: 31161530]
69. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon AR. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
70. Cheung DSK, Ho LYW, Chan LCK, Kwok RKH, Lai CKY. A home-based dyadic music-with-movement intervention for people with dementia and caregivers: a hybrid type 2 cluster-randomized effectiveness-implementation design. Clin Interv Aging 2022;17:1199-1216 [FREE Full text] [doi: 10.2147/CIA.S370661] [Medline: 35978943]

Abbreviations

ADL: activities of daily living
ART: Attention Restoration Theory
BPSD: behavioral and psychological symptoms of dementia
F-T-F: face-to-face
HT: horticultural therapy
ICC: intraclass correlation coefficient
MoCA-5-min: Montreal Cognitive Assessment-5-Minute
NGO: nongovernmental organization
NPI-D: Neuropsychiatric Inventory-Distress Scale
NPI-Q: Neuropsychiatric Inventory Questionnaire
RCT: randomized controlled trial
SHS: Subjective Happiness Scale
SRT: Stress Recovery Theory

Edited by X Xiang; submitted 02.09.24; peer-reviewed by A O'Donnell, LC Lu; comments to author 27.11.24; revised version received 05.12.24; accepted 18.12.24; published 05.02.25.

Please cite as:
Kor PPK, Liu JYW, Wong AKC, Tsang APL, Tan HZ, Cheung DSK, Leung HKW, Wong FKY
Effectiveness of a Dyadic Technology–Enhanced Home-Based Horticultural Therapy on Psychosocial Well-Being Among People With Dementia and Their Family Caregivers: Multimethods Pilot Study
JMIR Aging 2025;8:e66017
URL: https://aging.jmir.org/2025/1/e66017
doi:10.2196/66017
PMID:

Original Paper

Enhancing Older Adults' Lives Through Positive Aging Perception, Quality-of-Life Enhancement, and Social Support to Drive Acceptance and Readiness Toward Indoor Assistive Technology: Cross-Sectional Study

Ka Po Wong[1], PhD; Pei-Lee Teh[2], PhD; Weng Marc Lim[3], PhD; Shaun Wen Huey Lee[4], PhD

[1]Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Gerontechnology Laboratory and School of Business, Monash University Malaysia, Sunway City, Malaysia

[3]Sunway Business School, Sunway University, Sunway City, Malaysia

[4]School of Pharmacy, Monash University Malaysia, Sunway City, Malaysia

Corresponding Author:
Pei-Lee Teh, PhD
Gerontechnology Laboratory and School of Business
Monash University Malaysia
Jalan Lagoon Selatan
Sunway City, 47500
Malaysia
Phone: 60 355144971
Email: teh.pei.lee@monash.edu

Abstract

Background: The growing aging population faces increasing mobility limitations, highlighting the need for assistive technologies as potential solutions. These technologies support the independence and well-being of older adults and individuals with mobility challenges. Indoor mobility is essential for daily activities and significantly impacts their lives. Limited indoor mobility can reduce quality of life and heighten the risk of falls.

Objective: This study explores how positive aging perceptions, quality-of-life enhancements, and social support influence the acceptance and readiness of indoor assistive technologies among older adults.

Methods: A cross-sectional study was conducted at a gerontechnology laboratory, requiring participants to visit the facility in person. Each 60-minute session included demonstrations of various indoor assistive technologies and the completion of a questionnaire. The assistive technologies showcased encompassed a wide range of devices. Participants' positive aging perceptions, quality-of-life enhancements, social support, technology acceptance, and readiness were measured using validated scales. Data were analyzed with AMOS (version 28; IBM Corp) and SPSS (version 28; IBM Corp), using structural equation modeling and multivariate analysis of covariance to assess the effects of predictors while controlling for demographic factors.

Results: A total of 104 older adults aged 60 years and older participated, with a mean age of 67.92 (SD 5.68) years. Structural equation modeling indicated that positive aging perception has a significant influence on older adults' control beliefs (P=.095), comfort (P=.047), and confidence (P<.001) in gerontechnology. Multivariate analysis revealed significant combined effects of quality-of-life enhancement (P=.01) and social support (P=.03) on technology acceptance and readiness, wherein quality-of-life enhancement (P=.001) and social support (P=.008) negatively impacted security perception. Among demographic variables, educational level significantly impacted gerontechnology confidence (P=.004) while ethnicity influenced optimism (P=.003).

Conclusions: This study sheds light on key factors affecting older adults' acceptance and readiness to adopt indoor assistive technologies. Findings highlight the importance of fostering positive aging perceptions through these technologies. Addressing issues related to control beliefs, comfort, and confidence in gerontechnology is essential to enhance technology acceptance and readiness among older adults. Future research should investigate the underlying mechanisms and create targeted interventions to support successful technology adoption in this population.

(JMIR Aging 2025;8:e59665) doi:10.2196/59665

KEYWORDS

indoor assistive technology; positive aging perceptions; quality of life; social support; technology acceptance; technology readiness

Introduction

Background

Recent years have seen an increased focus on addressing the needs of older adults and individuals with mobility challenges [1]. As the global population ages, the prevalence of mobility challenges has become a significant societal concern [2]. The World Health Organization (WHO) projects that by 2050, the number of people aged 60 years and older will reach 2 billion, with a substantial portion experiencing mobility limitations [3]. This demographic shift has intensified efforts to develop innovative solutions that support the independence and well-being of older adults and those with mobility challenges [4].

Indoor mobility, which includes activities such as sitting, standing, and walking, is essential for maintaining independence and overall well-being, especially among older adults and individuals with mobility difficulties [5,6]. Limited indoor mobility can lead to serious consequences, including an increased risk of falls and a diminished quality of life [5]. Therefore, enhancing indoor mobility is crucial for enabling these individuals to participate actively in daily activities and sustain their independence [7]. Assistive technologies, such as commodes, home care beds, and reclining wheelchairs, have emerged as solutions to address the specific needs of this population. These technologies help overcome mobility limitations and improve functional abilities [8], potentially transforming how older adults and individuals with mobility challenges interact with their living environments. Successful integration of assistive technologies requires careful selection and personalization to match the unique capabilities, needs, and preferences of each user [9].

Indoor assistive technologies encompass a wide range of tools, devices, and equipment designed to enhance and support the independence and mobility of older adults and individuals with functional limitations within indoor settings [10]. These technologies include mobility aids (eg, scooters, walkers, and wheelchairs), transfer and positioning devices (eg, grab bars, patient lifts, and reclining chairs), smart home systems (eg, automated lighting and voice-controlled appliances), and various adaptive equipment (eg, bed rails, shower chairs, and toilet risers). Tailored to address the specific challenges faced by older adults and those with mobility challenges, these technologies facilitate daily activities and help maintain independence at home and in other indoor environments [11]. Through targeted support to increase functional capacity, indoor assistive technologies empower older adults and individuals with mobility challenges, enhancing their daily lives [12] and improving overall quality of life and well-being [13,14].

Gerontechnology, the intersection of gerontology and advanced technology, aims to enhance the health, independence, and quality of life of older adults [15]. Incorporating gerontechnological advancements into indoor assistive technologies addresses the unique challenges faced by this population, fostering greater independence and well-being [15]. Examples of gerontechnology applications include advanced mobility aids such as smart walkers and wheelchairs, cognitive support tools such as memory aids and smart home systems, and social engagement platforms such as web-based communication tools [16]. These technologies assist older adults in maintaining physical movement, performing daily activities, and reducing social isolation, thereby creating an environment where they can live independently and with dignity. Acceptance and readiness to adopt these technologies are influenced by positive aging perceptions, enhanced quality of life, and social support systems. This study examines how these factors drive technology acceptance and readiness among older adults, ultimately improving their lives through indoor assistive technologies.

Theoretical Background

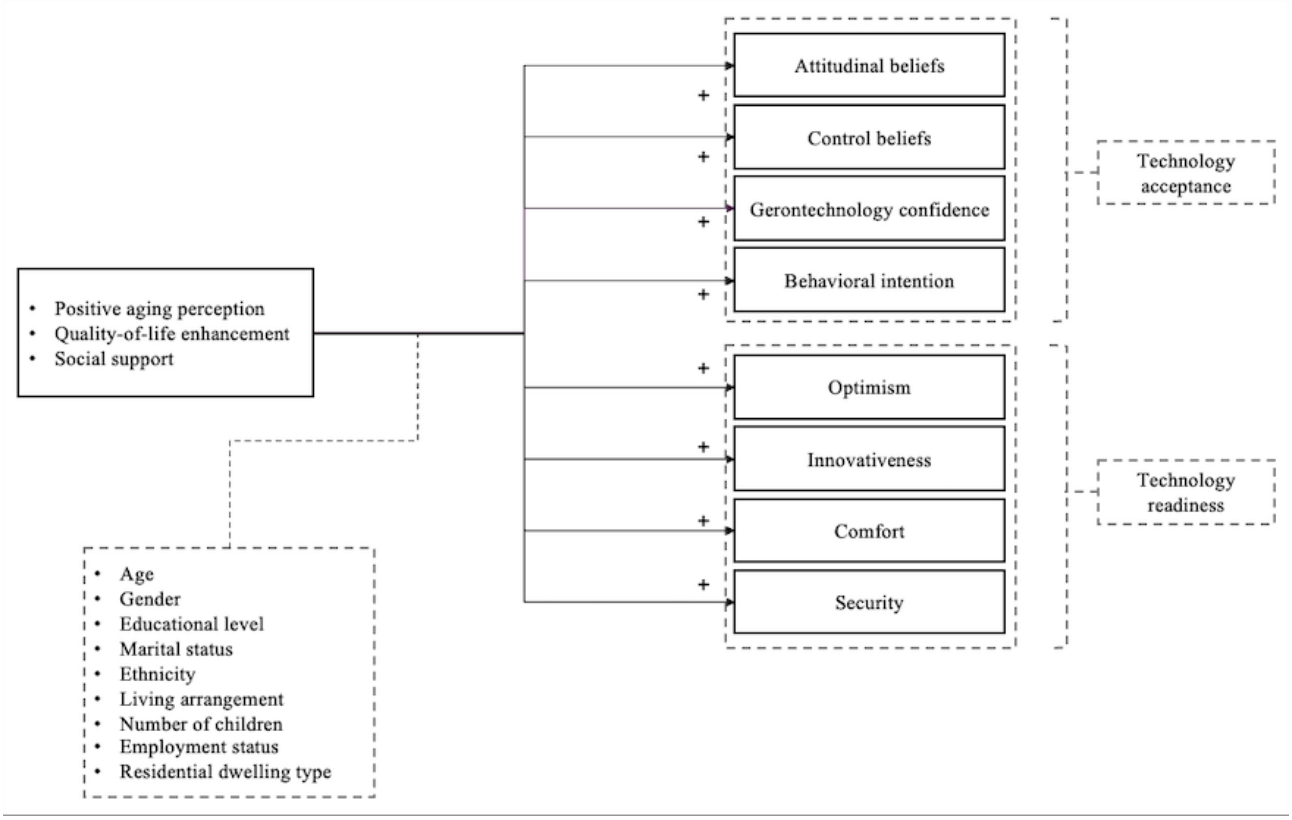
This study is grounded in the Theory of Planned Behavior (TPB) [17]. The TPB posits that individuals' behavioral intentions are influenced by 3 primary constructs: attitudes, subjective norms, and perceived behavioral control. In the context of this study, attitudes represent older adults' overall evaluation and perception of indoor assistive technologies. This study investigates whether older adults who hold a positive aging perception and believe that assistive technologies enhance their quality of life are more likely to adopt and use these technologies [18,19].

Subjective norms within the TPB refer to the social influences and societal expectations that shape an individual's behavior [20]. In this study, subjective norms are assessed through the role of social support—specifically, the assistance provided by family and friends—in influencing the acceptance and readiness to adopt indoor assistive technologies among older adults.

Perceived behavioral control is another critical construct of the TPB, encompassing individuals' beliefs and perceptions about their ability to perform a specific behavior [21,22]. This study measures perceived behavioral control by evaluating older adults' perceptions of their confidence and readiness to adopt and use indoor assistive technologies.

Applying the TPB enables this study to identify the factors that drive various aspects of older adults' acceptance and readiness to embrace indoor assistive technologies. Such insights are valuable for developing interventions and strategies aimed at promoting engagement, well-being, and overall quality of life among older adults [23]. The study's conceptual model is illustrated in Figure 1.

Figure 1. Study conceptual model.



Recognizing the needs of older adults with diminished mobility highlights the importance of creating innovative strategies to enhance their independence and physical and psychological well-being [1]. Although existing studies have examined the perceptions of technologies among older adults with and without disabilities or impairments, there remains a need for more evidence-based research focused on the specific factors that influence the acceptance and readiness to adopt indoor assistive technologies. Understanding these factors can provide promising pathways for improving indoor mobility and enhancing the quality of life for this population [13,22].

Previous research has explored various dimensions of indoor assistive technologies. For example, Gitlin et al [24] discovered that assistive devices and home modifications significantly improved functional abilities and reduced the risk of falls among older adults, indicating positive attitudes toward these technologies. Similarly, several studies have shown that the use of assistive devices in home environments leads to increased independence and reduced caregiver burden [15,25], underscoring the role of social support in technology acceptance. In addition, Demiris et al [26] and Liu et al [27] examined older adults' perceptions of their ability to use smart home technologies, aligning with the concept of perceived behavioral control in the TPB [17]. More recent studies, such as Peek et al [19], have identified key factors influencing the acceptance of technology for aging in place, including perceived ease of use, usefulness, and social influence, while Mitzner et al [28] highlighted the importance of user experience in technology acceptance among older adults. These studies provide empirical evidence supporting the benefits of indoor assistive technologies and emphasize the necessity of understanding user acceptance and readiness.

The global demographic shift toward an aging population and the rising prevalence of mobility limitations present urgent challenges that must be addressed. With the projected increase in the number of older adults, developing interventions and strategies that effectively cater to their unique needs and preferences becomes crucial [4]. To achieve a comprehensive understanding of older adults' needs regarding assistive technologies, targeted strategies must investigate their perspectives, acceptance, and readiness of aging in relation to indoor assistive technologies. Conducting such research can bridge existing knowledge gaps and inform the creation of strategies tailored to the specific needs of older adults, ultimately enhancing their overall independence and well-being [5].

This study aims to explore the influences of positive aging perception, quality-of-life enhancement through assistive technologies, and social support on technology acceptance and readiness among older adults in an indoor setting. Specifically, this study seeks to understand how these factors interact and contribute to older adults' attitudes and readiness to adopt and use indoor assistive technologies. The hypothesis posits that a positive aging perception, improved quality of life, and supportive social relationships positively affect older adults' technology acceptance, including attitudinal beliefs, control beliefs, confidence in gerontechnology, and behavioral intention. It is further hypothesized that these factors also positively impact older adults' readiness, which encompasses comfort, innovativeness, optimism, and security.

Methods

Study Setting

This study used a cross-sectional experiential design conducted at a gerontechnology laboratory located at an Australian university's international branch campus in Malaysia. Participants were required to visit the laboratory in person to take part in the study. Each session lasted approximately 60 minutes and included active participation in demonstrations of various indoor assistive technologies, followed by the completion of a questionnaire. The assistive technology demonstrations encompassed a wide range of devices tailored to support mobility.

Participants Recruitment

Recruitment efforts targeted individuals through advertisements placed across multiple social media platforms and in universities. Interested participants registered their interest by completing an online registration form through Google Forms or by contacting the provided phone number. Eligibility criteria were established to ensure suitable participation, including being aged 60 years or older; not having severe mobility challenges that would prevent full engagement in the experiential session; the ability to attend the demonstration session; and the capacity to provide written informed consent. These criteria ensured that participants could actively engage in the study activities and provide meaningful feedback on the assistive technologies demonstrated.

Study Design

Participants engaged in an experiential session featuring various indoor assistive technologies (given in Figure 2A-J). Each demonstrated device is specifically designed to support older adults and individuals with mobility challenges within their living environments [29]. The session included the following technologies: The SafeFree handling side guard and home care bed (Figure 2A) provide safe and comfortable support for resting, sleeping, and transferring activities within the home. The transfer roller kit board (Figure 2B) assists with safe transfers from beds, chairs, or the floor, enhancing indoor mobility. The tilt and reclining wheelchair (Figure 2C) allows users to adjust their position and posture to perform different activities at home. The shower commode chair (Figure 2D) enables individuals to bathe and use the toilet safely within their living spaces. The commode with a mobile pan (Figure 2E) offers discreet and convenient toileting support indoors. The recliner and stand-up sofa (Figure 2F) incorporate features that aid in sitting, standing, and repositioning for those with mobility challenges. The antislip seating mat (Figure 2G) enhances stability and safety on various seating surfaces within the home. The carbon fiber quad cane (Figure 2H) provides additional support and stability for individuals with limited mobility. The 3-in-1 stand assist walker (Figure 2I) supports users in performing various daily activities safely indoors. Finally, the nonintrusive activity-based sensor monitoring system (Figure 2J) offers remote monitoring and support for individuals in their living spaces.

Figure 2. Assistive technologies demonstrated in the experiential session: (A) handling side guard and home care bed; (B) transfer roller kit board; (C) tilt and reclining wheelchair; (D) shower commode chair; (E) commode with mobile pan; (F) recliner and stand-up sofa; (G) antislip seating mat; (H) carbon fiber quad cane; (I) 3-in-1 stand assist walker; and (J) nonintrusive activity-based sensor monitoring system.

A trained research assistant conducted the demonstrations to ensure consistency and quality throughout the session. Each participant received personalized attention from a single research assistant to maintain fidelity. The research team reviewed and approved all demonstration materials to ensure accuracy and effectiveness. Before the study, the research assistant conducted a trial demonstration for the research team, evaluating the clarity and conveyance of the content. In addition, a research team member conducted random visits to demonstration sessions to verify that the content remained consistent and adhered to the predetermined standards.

Outcome Measures

Participants were assessed after the demonstration session using a questionnaire divided into 5 main sections: positive aging

perception, quality-of-life enhancement, social support, technology acceptance, and technology readiness.

Positive Aging Perception

The Awareness of Age-Related Change (AARC) questionnaire [30] was used to assess participants' positive aging perception. A total of 10 items were selected for this study, using a 5-point Likert scale ranging from 1 (not at all) to 5 (very much). The internal consistency reliability demonstrated strong results for both AARC-Gains (α=.82) and AARC-Losses (α=.78) [31].

Quality-of-Life Enhancement

Quality-of-life enhancement was measured based on Moxley et al [32]. This section comprised 5 items evaluated using a 7-point Likert scale, from 1 (Not at all) to 7 (A lot). The items specifically addressed the extent to which assistive technologies can improve the quality of life for older adults. This scale has been widely adopted in gerontechnology research including studies by Dale et al [33].

Social Support

Social support was assessed using a 5-item scale adapted from Moxley et al [32]. This scale evaluated the level of assistance required from family and friends in becoming proficient with the presented technologies. Responses were measured on a 7-point Likert scale, ranging from 1 (Not at all) to 7 (A lot). The scale's widespread use in gerontechnology research, as demonstrated by its application in Dale et al [33], underscores its validity and reliability.

Technology Acceptance

Technology acceptance was evaluated using the Senior Technology Acceptance Model questionnaire, adapted from Chen and Lou [34] and Venkatesh et al [35]. This instrument included 4 variables: attitudinal beliefs (3 items), control beliefs (4 items), gerontechnology confidence (gerontechnology anxiety; 2 items), and behavioral intention (3 items). Participants responded on a 7-point Likert scale, from 1 (Strongly disagree) to 7 (Strongly agree). The questionnaire demonstrated strong internal consistency, with α values ranging from 0.96 to 0.97 [36]. Its widespread adoption across various fields highlights its effectiveness in assessing users' perceptions and attitudes toward technology adoption and utilization [37].

Technology Readiness

The Technology Readiness Questionnaire (TRQ) developed by Parasuraman and Colby [38] was used to measure participants' readiness and willingness to embrace technology. The TRQ includes 4 dimensions: comfort, innovativeness, optimism, and security, with 4 items assigned to each dimension. Participants rated each item on a 5-point Likert scale, where higher scores indicated a greater inclination to adopt and use technology. The TRQ's extensive application in aging-related research emphasizes its relevance in evaluating individuals' attitudes and preparedness for technology adoption [39].

Analytical Techniques

The measurement model was validated using confirmatory factor analysis (CFA) with SPSS (version 28.0; IBM Corp) and AMOS (version 28; IBM Corp). CFA evaluated the validity of the measurement model by examining the relationships between observed variables and their underlying latent constructs. A total of 8 key goodness-of-fit indices, recommended by Hu and Bentler [40], were used to assess the model fit. These indices included Cronbach α, chi-square and its respective degrees of freedom (χ^2/df), goodness-of-fit index (GFI), incremental fit index (IFI), comparative fit index (CFI), Tucker-Lewis index (TLI), root mean square error of approximation (RMSEA), and parsimony goodness-of-fit index (PGFI).

Descriptive statistics summarized the demographic variables, providing the mean and SD for each measure to offer an overview of the sample characteristics. Correlation analysis was performed to provide a preliminary assessment of the relationships between variables, while structural equation modeling explored the main relationships between predictor and outcome variables. Multivariate analysis of covariance (MANCOVA) was used as a post hoc analysis to scrutinize the effects of predictor variables on outcome variables while controlling for demographic characteristics such as age, gender, education level, and ethnicity. Wilks lambda (λ) was used to assess multivariate effects, followed by univariate analyses to identify specific between-subjects effects. Effect sizes were reported using partial eta squared (η^2). Statistical significance was set at $P<.05$, with marginal significance noted at $P<.10$.

Ethical Considerations

The research protocol was developed by Monash University Malaysia, and study approval was granted by the Monash University Human Research Ethics Committee (project ID: 39857; review reference: 2023-39857-98651) in September 2023. All participants provided written informed consent and received a token of appreciation of RM 50 (approximately US $11.2) for their participation. To ensure participant confidentiality, all collected data will be anonymized. In instances where full anonymization is not possible, strict protective measures will be implemented, including secure data storage and restricted access to authorized personnel only, to safeguard participant information.

Results

Participant Characteristics

A total of 104 older adults participated in the study, with an average mean age of 67.92 (SD 5.68) years. The sample comprised 58.7% (61/104) women and 41.3% (43/104) men. Educational levels varied, including 2.9% (3/104) with primary education and below, 24% (25/104) with secondary education, 27.9% (29/104) holding a diploma or preuniversity qualification, 29.8% (31/104) possessing a degree or professional certification, and 15.4% (16/104) with postgraduate degrees. Most participants were married (73/104, 70.2%), followed by those who were never married (5.8%, 6/104), divorced (9/104, 8.7%), widowed or widowers (10/104, 9.6%), separated (5/104, 4.8%), and others (1/104, 1%). Ethnically, the majority identified as Chinese (86/104, 82.7%), with smaller representations of Indian (16/104, 15.4%), Malay (1/104, 1%), and other ethnicities (1/104, 1%). Regarding living arrangements, 78.8% (82/104) resided with household members, 17.3% (18/104) lived alone, and 3.8% (4/104) had other living situations. Participants reported having

varying numbers of children, with 14.4% (15/104) having none, 14.4% (15/104) having 1 child, 32.7% (34/104) having 2 children, 32.7% (34/104) having 3 children, 1.9% (2/104) having 4 children, and 2.9% (3/104) having 5 children. In terms of employment status, 57.7% (60/104) were private retirees, 11.5% (12/104) were government retirees (including pensioners), 8.7% (9/104) were self-employed, 5.8% (6/104) were private sector employees, 7.7% (8/104) were homemakers, 5.8% (6/104) were unemployed, and 2.9% (3/104) were in other categories. The majority of participants lived in terraced, linked houses, or semidetached homes (60/104, 57.7%), followed by flats, apartments, condominiums, or townhouses (34/104, 32.7%); detached houses (9/104, 8.7%); and other dwelling types (1/104, 1%). Statistical analysis revealed that living arrangements significantly varied within the sample (P=.002), while other demographic variables did not show significant differences (Table 1).

Table 1. Demographic information (N=104).

Demographics	Values	*P* value
Age (years), mean (SD)	67.92 (5.68)	.22
Gender, n (%)		.08
Men	43 (41.3)	
Women	61 (58.7)	
Educational level, n (%)		.97
Primary school and below	3 (2.9)	
Secondary school	25 (24)	
Diploma or preuniversity	29 (27.9)	
Degree or professional	31 (29.8)	
Postgraduate	16 (15.4)	
Marital status, n (%)		.16
Never married	6 (5.8)	
Married	73 (70.2)	
Separated	5 (4.8)	
Divorcee	9 (8.7)	
Widow or widower	10 (9.6)	
Others	1 (1)	
Ethnicity, n (%)		.18
Malay	1(1)	
Chinese	86 (82.7)	
Indian	16 (15.4)	
Others	1 (1)	
Living arrangement, n (%)		.002
Living alone	18 (17.3)	
Living with household members	82 (78.8)	
Others	4 (3.8)	
Number of children, n (%)		.39
0	15 (14.4)	
1	15 (14.4)	
2	34 (32.7)	
3	34 (32.7)	
4	2 (1.9)	
5	3 (2.9)	
Employment status, n (%)		.41
Private sector employee	6 (5.8)	
Self-employed	9 (8.7)	
Government retirees (including pensioners)	12 (11.5)	
Private retiree	60 (57.7)	
Homemaker	8 (7.7)	
Unemployed	6 (5.8)	
Others	3 (2.9)	
Residential dwelling type, n (%)		.18

Demographics	Values	*P* value
Flat, apartment, condominium, or townhouse	34 (32.7)	
Detached house (bungalow or traditional house)	9 (8.7)	
Terrace, link house, or semidetached	60 (57.7)	
Others	1 (1)	

Correlations

The correlation analysis in Table 2 identified several significant relationships among the study variables. Age positively correlated with marital status (r= 0.278; P=.004), indicating that older participants were more likely to be married. Gender exhibited a significant negative correlation with living arrangements (r=–0.309; P=.001), suggesting that women were more likely to live with household members compared with men. Educational level was negatively associated with employment status (r=–0.233; P=.02). Residential dwelling type showed a positive correlation with living arrangements (r=0.397; P<.001) and the number of children (r=0.207; P=.04), indicating that certain housing types were more common among those living with more children. Quality-of-life enhancement was negatively related to age (r=–0.220; P=.02) and educational level (r=–0.201; P=.04). Social support positively correlated with quality-of-life enhancement (r=0.223; P=.02) and negatively correlated with gerontechnology confidence (r=–0.317; P=.001). Behavioral intention was positively associated with quality-of-life enhancement (r=0.246; P=.01), attitudinal beliefs (r=0.275; P=.005), control beliefs (r=0.396; P<.001), gerontechnology confidence (r=0.206; P=.04), optimism (r=0.428; P<.001), and innovativeness (r=0.204; P=.04). These correlations offer preliminary insights on how demographic factors and positive perceptions of aging, quality of life, and social support influence technology acceptance and readiness among older adults.

Table 2. Correlation matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Age (mean 67.92, SD 5.68)																				
r	1	–.063	–.131	.278	.112	–.135	.199	–.014	.019	–.020	–.220	.01	–.129	.195	.042	–.068	.063	.132	.089	–.034
P value	—[a]	.52	.18	.004	.26	.17	.04	.89	.84	.84	.02	.92	.19	.047	.67	.49	.52	.18	.37	.73
2. Gender[b] (mean 1.59, SD 0.495)																				
r	–.063	1	.003	.14	.131	–.309	–.085	.08	–.133	.085	–.143	.096	.025	.035	–.045	–.010	–.050	–.189	–.099	–.150
P value	.52	—	.97	.16	.18	.001	.39	.42	.18	.39	.15	.33	.80	.72	.65	.92	.62	.06	.32	.13
3. Educational level[c] (mean 3.31, SD 1.089)																				
r	–.131	.003	1	–.162	–.140	–.045	–.137	–.233	.059	.178	–.201	–.289	.125	.114	.406	.088	.011	.025	.051	.095
P value	.18	.97	—	.101	.16	.65	.17	.02	.55	.07	.04	.003	.21	.25	<.001	.38	.91	.80	.61	.34
4. Marital status[d] (mean 2.49, SD 1.115)																				
r	.278	.14	–.162	1	.042	–.089	.107	.117	.007	.159	–.034	.025	.103	.169	.025	.008	–.045	.277	–.002	.048
P value	.004	.16	.101	—	.67	.37	.28	.24	.95	.11	.73	.80	.30	.09	.80	.94	.65	.004	.98	.63
5. Ethnicity[e] (mean 2.19, SD 0.609)																				
r	.112	.131	–.140	.042	1	.125	.046	.019	.128	–.043	.269	.260	.054	.146	–.030	.214	.268	.043	.098	–.092
P value	.26	.18	.16	.67	—	.21	.64	.85	.19	.66	.006	.008	.58	.14	.76	.03	.006	.66	.32	.36
6. Living arrangement[f] (mean 1.9, SD 0.566)																				
r	–.135	–.309	–.045	–.089	.125	1	.129	–.002	.397	–.060	.079	–.078	.065	.02	.182	.07	.16	.024	.195	.244
P value	.17	.001	.65	.37	.21	—	.19	.99	<.001	.55	.42	.43	.51	.84	.06	.48	.11	.81	.047	.01
7. Number of children (mean 2.02, SD 1.188)																				
r	.199	–.085	–.137	.107	.046	.129	1	–.111	.207	.043	–.001	.026	–.135	.052	.004	.011	.085	.031	–.077	–.055
P value	.04	.39	.17	.28	.64	.19	—	.26	.04	.66	.99	.80	.17	.60	.97	.91	.39	.75	.44	.58
8. Employment status[g] (mean 5.67, SD 1.517)																				
r	–.014	.08	–.233	.117	.019	–.002	–.111	1	.01	–.017	.049	.002	–.105	–.108	.01	–.062	–.075	.152	.067	.031
P value	.89	.42	.02	.24	.85	.99	.26	—	.92	.86	.62	.98	.29	.28	.92	.53	.45	.12	.50	.75
9. Residential dwelling type[h] (mean 2.28, SD 0.96)																				
r	.019	–.133	.059	.007	.128	.397	.207	.01	1	–.016	–.010	–.162	–.064	–.088	.093	.036	.105	.034	.104	.179
P value	.84	.18	.55	.95	.19	<.001	.04	.92	—	.87	.92	.101	.52	.38	.35	.71	.29	.73	.29	.07
10. Positive ageing perception (mean 3.959, SD 0.404)																				
r	–.020	.085	.178	.159	–.043	–.060	.043	–.017	–.016	1	–.077	–.090	.192	.053	.202	–.082	–.008	.073	.094	–.008
P value	.84	.39	.07	.11	.66	.55	.66	.86	.87	—	.44	.36	.051	.59	.04	.41	.93	.46	.34	.94
11. Quality-of-life enhancement (mean 4.687, SD 1.356)																				
r	–.220	–.143	–.201	–.034	.269	.079	–.001	.049	–.010	–.077	1	.223	.082	.175	–.061	.246	.201	.161	.072	.19
P value	.02	.15	.04	.73	.006	.42	.99	.62	.92	.44	—	.02	.41	.07	.54	.01	.04	.10	.47	.054
12. Social support (mean 3.483, SD 1.295)																				

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
r	.01	.096	–.289	.025	.260	–.078	.026	.002	–.162	–.090	.223	1	.16	.041	–.317	–.093	–.123	–.109	–.150	–.232
P value	.92	.33	.003	.80	.008	.43	.80	.98	.101	.36	.02	—	.104	.68	.001	.35	.21	.27	.13	.02
13. Attitudinal beliefs (mean 5.401, SD 1.293)																				
r	–.129	.025	.125	.103	.054	.065	–.135	–.105	–.064	.192	.082	.16	1	.276	.177	.275	.281	.129	.102	.152
P value	.19	.80	.21	.30	.58	.51	.17	.29	.52	.051	.41	.104	—	.005	.07	.005	.004	.19	.30	.12
14. Control beliefs (mean 5.337, SD 0.931)																				
r	.195	.035	.114	.169	.146	.02	.052	–.108	–.088	.053	.175	.041	.276	1	.201	.396	.177	.341	.022	.1
P value	.047	.72	.25	.09	.14	.84	.60	.28	.38	.59	.08	.68	.005	—	.04	<.001	.07	<.001	.82	.31
15. Gerontechnology confidence (mean 4.88, SD 1.471)																				
r	.042	–.045	.406	.025	–.080	.182	.004	.01	.093	.202	–.061	–.317	.177	.201	1	.206	.157	.003	.408	.335
P value	.67	.65	<.001	.80	.76	.06	.97	.92	.35	.04	.54	.001	.07	.04	—	.04	.11	.98	<.001	.001
16. Behavioral intention (mean 5.804, SD 0.786)																				
r	–.068	–.010	.088	.008	.214	.07	.011	–.062	.036	–.082	.246	–.093	.275	.396	.206	1	.428	.204	.076	.07
P value	.49	.92	.38	.94	.03	.48	.91	.53	.71	.41	.01	.35	.005	<.001	.04	—	<.001	.04	.44	.48
17. Optimism (mean 4.276, SD 0.515)																				
r	.063	–.050	.011	–.045	.268	.16	.085	–.075	.105	–.008	.201	–.123	.281	.177	.157	.428	1	.269	.145	.103
P value	.52	.62	.91	.65	.006	.11	.39	.45	.29	.93	.04	.21	.004	.07	.11	<.001	—	.006	.14	.30
18. Innovativeness (mean 3.26, SD 0.637)																				
r	.132	–.189	.025	.277	.043	.024	.031	.152	.034	.073	.161	–.109	.129	.341	.003	.204	.269	1	.173	.053
P value	.18	.06	.80	.004	.66	.81	.75	.12	.73	.46	.102	.27	.19	<.001	.97	.04	.006	—	.08	.59
19. Comfort (mean 3.202, SD 0.626)																				
r	.089	–.099	.051	–.002	.098	.195	–.077	.067	.104	.094	.072	–.150	.102	.022	.408	.076	.145	.173	1	.404
P value	.37	.32	.61	.98	.32	.047	.44	.50	.29	.34	.47	.13	.30	.82	<.001	.44	.14	.08	—	<.001
20. Security (mean 2.752, SD 0.746)																				
r	–.034	–.150	.095	.048	–.092	.244	–.065	.031	.179	–.008	.19	–.232	.152	.1	.335	.07	.103	.053	.404	1
P value	.73	.13	.34	.63	.36	.01	.58	.75	.07	.94	.054	.02	.12	.31	.001	.48	.30	.59	<.001	—

[a]Not applicable.

[b]Gender: 1=men and 2=women.

[c]Educational level: 1=primary school and below, 2=secondary school, 3=diploma or preuniversity, 4=degree or professional, and 5=postgraduate.

[d]Marital status: 1=never married, 2=married, 3=separated, 4=divorcee, 5=widow or widower, and 6=others.

[e]Ethnicity: 1=Malay, 2=Chinese, 3=Indian, 4=others.

[f]Living arrangement: 1=living alone, 2=living with household members, and 3=others.

[g]Employment status: 1=private sector employee, 2=self-employed, 3=government retiree (includes pensioners), 4=private retiree, 5=homemaker, 6=unemployed, and 7=others.

[h]Residential dwelling type: 1=flat, apartment, condominium, or townhouse; 2=detached house (bungalow or traditional house); 3=terrace, link house, or semidetached; 4=others.

Measurement Model

The measurement model was validated via CFA. The Cronbach α values for positive aging perception, quality-of-life enhancement, social support, technology acceptance, and technology readiness were 0.736, 0.855, 0.815, 0.740, and 0.731, respectively, demonstrating acceptable internal consistency.

Model fit indices were as follows: χ^2_{752}=1069.140, *P*<.001; GFI=0.703; IFI=0.858; CFI=0.850; TLI=0.829; RMSEA=0.064; and PGFI=0.585. These values met the recommended thresholds, indicating a moderately acceptable model fit [40].

Structural Model

The structural model was tested via structural equation modeling (Table 3), illustrating how positive aging perception, quality-of-life enhancement, and social support relate to various aspects of technology acceptance and readiness among older adults. Positive aging perception significantly enhanced gerontechnology confidence (β=.462; *P*<.001) and comfort (β=.323; *P*=.047) and exhibited a marginally significant positive relationship with control beliefs (β=.228; *P*=.095). In contrast, quality-of-life enhancement and social support did not show significant effects on any of the technology acceptance or readiness components. These findings suggest that positive aging perception plays a crucial role in fostering confidence and comfort with gerontechnology, whereas quality-of-life enhancement and social support have limited impacts on technology acceptance and readiness within this population.

Table 3. Structural equation modeling results.

Path	Standardized coefficient (β)	Unstandardized coefficient (B)	SE	*t* value	*P* value
Positive aging perception					
Attitudinal beliefs	.143	0.173	0.145	1.196	.23
Control beliefs	.228	0.242	0.145	1.670	.095
Gerontechnology confidence	.462	1.220	0.321	3.806	<.001
Behavioral intention	.190	0.239	0.146	1.635	.102
Optimism	.041	0.027	0.068	0.400	.69
Innovativeness	.061	0.042	0.094	0.453	.65
Comfort	.323	0.227	0.114	1.984	.047
Security	.171	0.082	0.067	1.219	.22
Quality-of-life enhancement					
Attitudinal beliefs	.097	0.932	1.687	0.552	.58
Control beliefs	.031	0.262	1.012	0.259	.80
Gerontechnology confidence	–.104	–2.177	3.615	–0.602	.55
Behavioral intention	.266	2.652	3.973	0.667	.50
Optimism	.165	0.878	1.421	0.618	.54
Innovativeness	.149	0.829	1.382	0.600	.55
Comfort	.160	0.896	1.475	0.607	.54
Security	.150	0.572	0.953	0.600	.55
Social support					
Attitudinal beliefs	.186	0.484	0.398	1.217	.22
Control beliefs	.142	0.323	0.323	1.002	.32
Gerontechnology confidence	–.215	–1.220	0.858	–1.422	.16
Behavioral intention	–.124	–0.333	0.336	–.993	.32
Optimism	–.192	–0.277	0.245	–1.130	.26
Innovativeness	–.176	–0.264	0.242	–1.090	.28
Comfort	–.289	–0.436	0.329	–1.323	.19
Security	–.357	–0.367	0.279	–1.314	.19

Post hoc Observations

The MANCOVA assessed the combined effects of independent variables and demographic factors on technology acceptance and readiness among older adults (Tables 4 and 5). The analysis revealed that positive aging perception did not significantly influence the various aspects of technology acceptance and readiness when controlling for demographic characteristics (Wilks λ=0.897, $F_{8,61}$=0.877, *P*=.54, partial η^2=0.103). This consistency reinforces the structural model findings, indicating that the relationships between positive aging perception and factors such as control beliefs, comfort, and gerontechnology confidence remain stable across different demographic groups. Conversely, quality-of-life enhancement (Wilks λ=0.737, $F_{8,61}$=2.717, *P*=.01, partial η^2=0.263) and social support (Wilks

λ=0.762, $F_{8,61}$=2.387, *P*=.03, partial η²=0.238) significantly impacted technology acceptance and readiness. A closer examination showed that both quality-of-life enhancement (*P*=.001) and social support (*P*=.008) negatively influenced security perceptions, suggesting that improvements in these areas are associated with reduced concerns about the security of assistive technologies. In addition, demographic factors played a notable role: educational level significantly predicted gerontechnology confidence (β=29.548, *P*=.004, partial η²=0.201) and ethnicity was a significant predictor of optimism (β=3.373, *P*=.003, partial η²=0.187). These findings indicate that while positive aging perceptions consistently affect certain aspects of technology acceptance, quality-of-life enhancements and social support are crucial in shaping specific perceptions such as security. Furthermore, educational attainment and ethnic background influence confidence and optimism regarding technology use. These granular insights highlight the importance of addressing both demographic and psychological factors to foster effective adoption and readiness for indoor assistive technologies among older adults. Further explanation is provided in the Discussion section.

Table 4. Multivariate analysis of covariance on the effects of predictor variables while controlling for demographic characteristics.

Variables	Wilks λ	*F* value (*df*)	*P* value	Partial η²
Predictor variables				
Positive aging perception	0.897	0.877 (8, 61)	.54	0.103
Quality-of-life enhancement	0.737	2.717 (8, 61)	.01	0.263
Social support	0.762	2.387 (8, 61)	.03	0.238
Demographic characteristics				
Age	0.770	1.066 (8, 61)	.39	0.123
Gender	0.920	0.663 (8, 61)	.72	0.080
Educational level	0.498	1.471 (8, 61)	.06	0.160
Marital status	0.584	0.883 (8, 61)	.67	0.102
Ethnicity	0.571	1.574 (8, 61)	.051	0.170
Living arrangement	0.791	0.947 (8, 61)	.52	0.110
Number of children	0.642	0.718 (8, 61)	.90	0.085
Employment status	0.519	0.904 (8, 61)	.66	0.104
Residential dwelling type	0.753	0.759 (8, 61)	.78	0.090
Significant between-subjects path effects (*F* >3.00, *P* <.05)				
Quality-of-life enhancement: security	6.404	12.913 (1, 68)	.001	0.160
Social support: security	3.674	7.408 (1, 68)	.008	0.098
Educational level: gerontechnology confidence	29.548	4.266 (4, 68)	.004	0.201
Ethnicity: optimism	3.373	5.198 (3, 68)	.003	0.187

Table 5. Multiple regression analyses examining the effects of predictor variables and demographic characteristics on technology acceptance and technology readiness outcomes.

Outcome variables	R^2	Adjusted R^2	*F* value (*df*)	*P* value
Overall model effects (predictor variables+demographic characteristics=outcomes)				
Attitudinal beliefs	0.439	0.158	1.562 (7, 95)	.06
Control beliefs	0.398	0.097	1.321 (7, 95)	.16
Gerontechnology confidence	0.471	0.207	1.783 (7, 95)	.02
Behavioral intention	0.303	–0.045	0.870 (7, 95)	.67
Optimism	0.441	0.161	1.576 (7, 95)	.06
Innovativeness	0.348	0.021	1.066 (7, 95)	.40
Comfort	0.337	0.005	1.015 (7, 95)	.47
Security	0.401	0.101	1.339 (7, 95)	.15

Discussion

Principal Findings

This study explored the relationships between positive aging perception, quality-of-life enhancement, social support, and various factors related to technology acceptance and readiness among older adults.

The preliminary analysis using correlation analysis revealed significant associations between demographic factors and key variables, such as age being positively related to marital status and negatively associated with quality-of-life enhancement and gerontechnology confidence; educational level being negatively correlated with living arrangements and gerontechnology confidence; and social support being positively correlated with quality-of-life enhancement. These insights suggest that the older adults who participated in this study predominantly consisted of married individuals with varying educational backgrounds, most of whom lived with household members and received substantial social support, which in turn influenced their perceptions and confidence regarding assistive technologies.

The primary analysis using structural equation modeling indicated that a positive perception of aging significantly enhances gerontechnology confidence and comfort, with a marginal positive relationship with control beliefs. These findings suggest that older adults who maintain a positive outlook on aging are more confident and comfortable with using gerontechnology, which aligns with existing literature emphasizing the role of positive aging attitudes in technology adoption [41,42]. However, quality-of-life enhancement and social support did not exhibit significant direct effects on most components of technology acceptance and readiness within the structural model. This contrasts with some previous studies that highlighted the importance of these factors in technology adoption [43,44], indicating that their influence may be more complex or mediated by other variables not captured in this model.

The post hoc analysis using MANCOVA further revealed that quality-of-life enhancement and social support significantly impact technology acceptance and readiness when controlling for demographic characteristics. Specifically, both quality-of-life enhancement and social support were found to negatively influence security perceptions, suggesting that as these factors increase, concerns about the security of assistive technologies decrease. This negative influence on security perceptions could be explained by demographic factors such as educational level and ethnicity. Educational level significantly predicted gerontechnology confidence, indicating that individuals with higher education may feel more confident in using technology, thereby reducing their security concerns. Similarly, ethnicity emerged as a significant predictor of optimism, suggesting that cultural or social backgrounds influence positive expectations toward technology, which can also alleviate security apprehensions. These findings highlight that while quality-of-life enhancements and social support do not directly influence the larger set of technology acceptance factors, they play a crucial role in alleviating security-related concerns, which are vital for the overall readiness to adopt assistive technologies.

Interestingly, social support showed negative associations with gerontechnology confidence and security perceptions in the structural model, although these were not statistically significant. This counterintuitive finding, if significant, may indicate that excessive reliance on social support could undermine individuals' confidence in using technology independently or heighten concerns about privacy and security. Such dynamics, which contrast past observations [45-48], warrant further investigation to understand the complex role of social support in technology adoption among older adults. One possible explanation is the influence of social norms, where older adults may experience expectations or pressures from their social networks that discourage full technological engagement [49]. An overreliance on assistance from others might also reduce the perceived need for technology, leading to decreased confidence and increased security concerns. Privacy and trust issues related to technology could also contribute to these negative relationships, as older adults may fear privacy breaches or distrust the reliability of assistive technologies [50,51].

The significant influence of demographic characteristics, particularly educational level and ethnicity, emphasizes the importance of considering sociodemographic factors in technology acceptance models. Higher educational levels were associated with greater confidence in using gerontechnology, likely due to better technological literacy and problem-solving skills acquired through education. Ethnic background influencing optimism suggests that cultural factors play a role in shaping positive attitudes toward technology adoption. These findings extend previous research [43,52] by quantifying the impact of these demographic factors and highlighting their specific effects on different aspects of technology acceptance.

Contrary to initial expectations, this study did not find significant associations between positive aging perception and several factors, including attitudinal beliefs, behavioral intention, innovativeness, optimism, and security, whereas quality-of-life enhancement and social support did not produce significant relationships. This may suggest that older adults' attitudes toward indoor assistive technologies may be more complex or mediated by other factors such as perceived need, technology complexity, or previous experience with similar devices that were not considered in this study [19,49]. While the lack of significant associations may seem surprising, it is important to acknowledge that the adoption and acceptance of assistive technologies among older adults are influenced by multifaceted factors [53]. The variables examined herein might be influenced by additional variables or interact with each other in ways that were not captured in this particular study. Future studies could explore other variables or contextual factors, such as cognitive abilities, physical health, technological literacy, or specific characteristics of the assistive technologies themselves, to develop more comprehensive models of technology acceptance among older adults [54].

Implications and Recommendations

This study underscores the importance of fostering positive aging perceptions to enhance gerontechnology confidence and

comfort among older adults. To achieve this, stakeholders should develop and implement educational programs and workshops that clearly demonstrate the benefits and practical applications of assistive technologies. These initiatives can help older adults build a more optimistic outlook on aging and increase their confidence in using these devices effectively.

Integrating quality-of-life enhancements and social support into the design and implementation of assistive technologies is crucial for addressing security concerns. Assistive devices should incorporate features that directly improve users' daily living experiences, such as intuitive user interfaces and robust safety mechanisms. These enhancements can help reduce apprehensions about the reliability and security of the technologies, making them more acceptable and trustworthy for older adults.

Demographic factors, including educational level and ethnicity, significantly influence technology confidence and optimism. To accommodate varying educational backgrounds, assistive technology solutions should offer clear instructions, accessible interfaces, and comprehensive training materials. Technologies should also be culturally tailored to respect and integrate the diverse cultural preferences of different ethnic groups. This cultural sensitivity can promote greater acceptance and positive attitudes toward the use of assistive technologies among ethnically diverse populations.

The negative associations observed between social support and gerontechnology confidence and security perceptions indicate that excessive reliance on social networks might undermine independent technology use and heighten security concerns. To address this, support systems should be designed to empower older adults to use assistive technologies independently. Training programs for caregivers and family members should focus on encouraging autonomous use of technology rather than providing constant assistance. Addressing privacy and trust issues through transparent data handling practices and robust security features should further help alleviate security concerns, thereby enhancing older adults' readiness to adopt assistive technologies.

To this end, enhancing technology acceptance and readiness among older adults requires a comprehensive approach that includes promoting positive aging attitudes, improving quality of life through targeted technology features, tailoring solutions to meet diverse educational and cultural backgrounds, and structuring social support to foster independence. Implementing these strategies should facilitate the effective adoption of indoor assistive technologies, ultimately improving the independence and well-being of older adults.

Limitations and Future Directions

This study presents several limitations that should be considered when interpreting the findings. First, the sample comprised older adults aged 60 years and older who participated in an indoor setting, predominantly Chinese and married, with most living with household members. While this demographic reflects typical patterns in urban Malaysian environments, it may limit the generalizability of the results to wider populations, including those in rural areas or living in assisted living facilities. Future research should use stratified sampling techniques to include a more diverse range of settings and demographics, enhancing the applicability of the findings across different contexts.

Second, the cross-sectional design of the study restricts the ability to establish causal relationships between variables. The associations identified provide a snapshot at a specific point in time but do not account for changes and developments over time. Longitudinal studies or experimental designs would be beneficial in examining the causal effects and temporal dynamics between positive aging perception, quality-of-life enhancement, social support, and technology acceptance and readiness.

Third, reliance on self-report measures to assess constructs such as positive aging perception, quality-of-life enhancement, social support, and technology acceptance introduces potential biases, including social desirability and recall bias. These biases may affect the accuracy of the data collected. Future studies could incorporate objective measures or combine self-report instruments with other assessment methods to achieve a more comprehensive understanding of these constructs.

Fourth, this study focused on specific variables—positive aging perception, quality-of-life enhancement, and social support—to explore their influence on technology acceptance and readiness. Other relevant factors, such as cognitive abilities, physical health, technological literacy, and specific characteristics of assistive technologies, were not included in the analysis. Incorporating a wider range of variables in future research would provide a more comprehensive view of the factors influencing older adults' technology acceptance and readiness. Using a multidimensional framework could capture the complex interactions among various factors more effectively.

Finally, the sample primarily consisted of Malaysian participants living in urban areas, which limits the diversity and generalizability of the findings to other cultural or demographic groups. Future studies should endeavor to include a more varied sample to examine potential cultural or contextual differences in the relationships between the variables. Using cross-cultural research designs or multisite studies can help identify and account for cultural and contextual variations, thereby enhancing the relevance and applicability of the findings across different populations.

Conclusion

Understanding the factors that influence older adults' acceptance and readiness to adopt indoor assistive technologies is essential for enhancing their independence and well-being. This study demonstrates that a positive perception of aging significantly increases gerontechnology confidence and comfort among older adults, while quality-of-life enhancements and social support play crucial roles in reducing security concerns related to technology use. Demographic factors, particularly educational level and ethnicity, also significantly influence confidence and optimism toward technology adoption. These insights highlight the need for targeted strategies to effectively promote the adoption of assistive technologies. Future research should explore the underlying mechanisms that drive these relationships and develop customized interventions to support successful technology integration among older adults.

Acknowledgments

This study was supported by the AGELESS research program funded by the Malaysian Ministry of Higher Education Long Term Research Grant Scheme (LRGS/1/2019/UM/01/1/1, LR005-2019, GOV-000047). We also extend our gratitude to Asian Integrated Medical Sdn Bhd for the loan of assistive technologies, which were used solely for demonstration purposes within the scope of this study.

Conflicts of Interest

None declared.

References

1. Freiberger E, Sieber CC, Kob R. Mobility in older community-dwelling persons: a narrative review. Front Physiol 2020;11:881 [FREE Full text] [doi: 10.3389/fphys.2020.00881] [Medline: 33041836]
2. Newman AB. The epidemiology and societal impact of aging-related functional limitations: a looming public health crisis. J Gerontol A Biol Sci Med Sci 2023;78(Suppl 1):4-7 [FREE Full text] [doi: 10.1093/gerona/glad021] [Medline: 37325965]
3. Leeson GW. The growth, ageing and urbanisation of our world. J Popul Ageing 2018;11:107-115. [doi: 10.1007/s12062-018-9225-7]
4. Moxham LJ, Stutchbury TK, Spinks G, de Vet E, Ikutegbe V, Traynor V, et al. Understanding the assistive technology needs of people over 55: the future of mobility aids. Australas J Ageing 2019;38(4):e127-e134. [doi: 10.1111/ajag.12648] [Medline: 30950196]
5. Brustio PR, Mulasso A, D'Emanuele S, Zia G, Feletti L, Del Signore S, et al. Indoor mobility, frailty, and disability in community-dwelling older adults: a mediation model. Int J Environ Res Public Health 2022;19(18):11386 [FREE Full text] [doi: 10.3390/ijerph191811386] [Medline: 36141659]
6. Kelsey JL, Berry SD, Procter-Gray E, Quach L, Nguyen USDT, Li W, et al. Indoor and outdoor falls in older adults are different: the maintenance of balance, independent living, intellect, and zest in the elderly of boston study. J Am Geriatr Soc 2010;58(11):2135-2141 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03062.x] [Medline: 20831726]
7. Hashidate H, Shimada H, Shiomi T, Shibata M, Sawada K, Sasamoto N. Measuring indoor life-space mobility at home in older adults with difficulty to perform outdoor activities. J Geriatr Phys Ther 2013;36(3):109-114. [doi: 10.1519/JPT.0b013e31826e7d33] [Medline: 22976813]
8. Carver J, Ganus A, Ivey JM, Plummer T, Eubank A. The impact of mobility assistive technology devices on participation for individuals with disabilities. Disabil Rehabil Assist Technol 2016;11(6):468-477. [doi: 10.3109/17483107.2015.1027295] [Medline: 25815679]
9. Kabacińska K, Vu K, Tam M, Edwards O, Miller WC, Robillard JM. "Functioning better is doing better": older adults' priorities for the evaluation of assistive technology. Assist Technol 2023;35(4):367-373 [FREE Full text] [doi: 10.1080/10400435.2022.2113180] [Medline: 35972791]
10. Cook AM, Polgar JM. Assistive Technologies: Principles and Practice. London, UK: Elsevier Health Sciences; 2019.
11. Fiorini L, de Mul M, Fabbricotti I, Limosani R, Vitanza A, D'Onofrio G, et al. Assistive robots to improve the independent living of older persons: results from a needs study. Disabil Rehabil Assist Technol 2021;16(1):92-102. [doi: 10.1080/17483107.2019.1642392] [Medline: 31329000]
12. Löfqvist C, Nygren C, Brandt ?, Oswald F, Iwarsson S. Use of mobility devices and changes over 12 months among very old people in five European countries. Aging Clin Exp Res 2007;19(6):497-505. [doi: 10.1007/BF03324737] [Medline: 18172373]
13. Agree EM, Freedman VA. A quality-of-life scale for assistive technology: results of a pilot study of aging and technology. Phys Ther 2011;91(12):1780-1788 [FREE Full text] [doi: 10.2522/ptj.20100375] [Medline: 22003159]
14. Matlabi H, Parker SG, McKee K. The contribution of home-based technology to older people's quality of life in extra care housing. BMC Geriatr 2011;11:68 [FREE Full text] [doi: 10.1186/1471-2318-11-68] [Medline: 22040111]
15. Marasinghe KM. Assistive technologies in reducing caregiver burden among informal caregivers of older adults: a systematic review. Disabil Rehabil Assist Technol 2016;11(5):353-360. [doi: 10.3109/17483107.2015.1087061] [Medline: 26371519]
16. Damant J, Knapp M, Watters S, Freddolino P, Ellis M, King D. The impact of ICT services on perceptions of the quality of life of older people. J Assist Technol 2013;7(1):21. [doi: 10.1108/17549451311313183]
17. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
18. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
19. Peek STM, Wouters EJM, van Hoof J, Luijkx KG, Boeije HR, Vrijhoef HJM. Factors influencing acceptance of technology for aging in place: a systematic review. Int J Med Inform 2014;83(4):235-248 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.004] [Medline: 24529817]

20. Wu J, Song S. Older adults' online shopping continuance intentions: applying the technology acceptance model and the theory of planned behavior. International Journal of Human–Computer Interaction 2020;37(10):938-948. [doi: 10.1080/10447318.2020.1861419]
21. Orellano-Colón EM, Harrison-Cruz S, López-Lugo E, Ramos-Peraza S, Meléndez-Ortiz A, Ortiz-Torres J, et al. Assistive technology self-management intervention for older Hispanics: a feasibility study. Disabil Rehabil Assist Technol 2020;15(8):862-870 [FREE Full text] [doi: 10.1080/17483107.2019.1621954] [Medline: 31172826]
22. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
23. Beyer AK, Wolff JK, Freiberger E, Wurm S. Are self-perceptions of ageing modifiable? Examination of an exercise programme with vs. without a self-perceptions of ageing-intervention for older adults. Psychol Health 2019;34(6):661-676. [doi: 10.1080/08870446.2018.1556273] [Medline: 30628483]
24. Gitlin LN, Winter L, Dennis MP, Corcoran M, Schinfeld S, Hauck WW. A randomized trial of a multicomponent home intervention to reduce functional difficulties in older adults. J Am Geriatr Soc 2006;54(5):809-816 [FREE Full text] [doi: 10.1111/j.1532-5415.2006.00703.x] [Medline: 16696748]
25. Beyer S, Perry J. Promoting independence through the use of assistive technology. Tizard Learn Disabil Rev 2013;18:179-185. [doi: 10.1108/tldr-03-2013-0028]
26. Demiris G, Hensel BK, Skubic M, Rantz M. Senior residents' perceived need of and preferences for "smart home" sensor technologies. Int J Technol Assess Health Care 2008;24(1):120-124. [doi: 10.1017/S0266462307080154] [Medline: 18218177]
27. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
28. Mitzner TL, Boron JB, Fausset CB, Adams AE, Charness N, Czaja SJ, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010;26(6):1710-1721 [FREE Full text] [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
29. Mann WC, Helal S. Smart Technology for Aging, Disability, and Independence: The State of the Science. Hoboken, NJ: Wiley; 2006.
30. Kaspar R, Gabrian M, Brothers A, Wahl HW, Diehl M. Measuring awareness of age-related change: development of a 10-item short form for use in large-scale surveys. Gerontologist 2019;59(3):e130-e140 [FREE Full text] [doi: 10.1093/geront/gnx213] [Medline: 29401221]
31. Nehrkorn-Bailey AM, Rodriguez D, Forsyth G, Braun B, Burke K, Diehl M. Change in views of aging, physical activity, and physical health over 8 weeks: results from a randomized study. J Aging Phys Act 2023;31(4):666-678 [FREE Full text] [doi: 10.1123/japa.2022-0133] [Medline: 36708712]
32. Moxley J, Sharit J, Czaja SJ. The factors influencing older adults' decisions surrounding adoption of technology: quantitative experimental study. JMIR Aging 2022;5(4):e39890 [FREE Full text] [doi: 10.2196/39890] [Medline: 36416885]
33. Dale J, Nanton V, Day T, Apenteng P, Bernstein CJ, Grason Smith G, et al. Uptake and use of care companion, a web-based information resource for supporting informal carers of older people: mixed methods study. JMIR Aging 2023;6:e41185 [FREE Full text] [doi: 10.2196/41185] [Medline: 37733406]
34. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging 2020;4(3):igaa016 [FREE Full text] [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
35. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
36. Goh TT, Dai X, Yang Y. Benchmarking ChatGPT for prototyping theories: experimental studies using the technology acceptance model. BenchCouncil Trans Benchmarks Stand Eval 2023;3(4):100153. [doi: 10.1016/j.tbench.2024.100153]
37. Kwan RYC, Yeung JWY, Lee JLC, Lou VWQ. The association of technology acceptance and physical activity on frailty in older adults during the COVID-19 pandemic period. Eur Rev Aging Phys Act 2023;20(1):24 [FREE Full text] [doi: 10.1186/s11556-023-00334-3] [Medline: 38114901]
38. Parasuraman A, Colby CL. An updated and streamlined technology readiness index: TRI 2.0. J Serv Res 2014;18(1):59-74. [doi: 10.1177/1094670514539730]
39. Kim S, Chow BC, Park S, Liu H. The usage of digital health technology among older adults in Hong Kong and the role of technology readiness and eHealth literacy: path analysis. J Med Internet Res 2023;25:e41915 [FREE Full text] [doi: 10.2196/41915] [Medline: 37043274]
40. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1-55 [FREE Full text] [doi: 10.1080/10705519909540118]
41. Lê Q, Nguyen HB, Barnett T. Smart homes for older people: positive aging in a digital world. Future Internet 2012;4(2):607-617. [doi: 10.3390/fi4020607]
42. Offermann J, Wilkowska W, Ziefle M. Interplay of perceptions of aging, care, and technology acceptance in older age. Int J Hum Comput 2022;39(5):1003-1015 [FREE Full text] [doi: 10.1080/10447318.2022.2050542]

43. Shin HR, Um SR, Yoon HJ, Choi EY, Shin WC, Lee HY, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023;25:e41935 [FREE Full text] [doi: 10.2196/41935] [Medline: 37036760]
44. Delbreil E, Zvobgo G. Wireless sensor technology in dementia care: caregiver perceptions, technology take ‐ up and business model innovation. EuroMed J Bus 2013;8(1):79-97. [doi: 10.1108/emjb-05-2013-0019]
45. Huang G, Oteng SA. Gerontechnology for better elderly care and life quality: a systematic literature review. Eur J Ageing 2023;20(1):27 [FREE Full text] [doi: 10.1007/s10433-023-00776-9] [Medline: 37347277]
46. Aggar C, Sorwar G, Seton C, Penman O, Ward A. Smart home technology to support older people's quality of life: a longitudinal pilot study. Int J Older People Nurs 2023;18(1):e12489 [FREE Full text] [doi: 10.1111/opn.12489] [Medline: 35785517]
47. Gitlin LN. Assistive technology in the home and community for older people: psychological and social considerations. In: Scherer MJ, editor. Assistive Technology: Matching Device and Consumer for Successful Rehabilitation. Washington, DC: American Psychological Association; 2002.
48. Wei W, Gong X, Li J, Tian K, Xing K. A study on community older people's willingness to use smart home-an extended technology acceptance model with intergenerational relationships. Front Public Health 2023;11:1139667 [FREE Full text] [doi: 10.3389/fpubh.2023.1139667] [Medline: 37346111]
49. Tsertsidis A, Kolkowska E, Hedström K. Factors influencing seniors' acceptance of technology for ageing in place in the post-implementation stage: a literature review. Int J Med Inform 2019;129:324-333 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.06.027] [Medline: 31445274]
50. Bradford DK, Kasteren YV, Zhang Q, Karunanithi M. Watching over me: positive, negative and neutral perceptions of in-home monitoring held by independent-living older residents in an Australian pilot study. Ageing Soc 2017;38(7):1377-1398. [doi: 10.1017/s0144686x1700006x]
51. Schomakers EM, Biermann H, Ziefle M. Understanding privacy and trust in smart home environments. Cham: Springer; 2020 Presented at: HCI for Cybersecurity, Privacy and Trust. HCII 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-50309-3_34]
52. Schroeder T, Dodds L, Georgiou A, Gewald H, Siette J. Older adults and new technology: mapping review of the factors associated with older adults' intention to adopt digital technologies. JMIR Aging 2023;6:e44564 [FREE Full text] [doi: 10.2196/44564] [Medline: 37191976]
53. Chen K. Why do older people love and hate assistive technology? an emotional experience perspective. Ergonomics 2020;63(12):1463-1474. [doi: 10.1080/00140139.2020.1808714] [Medline: 32780683]
54. Oh SS, Kim KA, Kim M, Oh J, Chu SH, Choi J. Measurement of digital literacy among older adults: systematic review. J Med Internet Res 2021;23(2):e26145 [FREE Full text] [doi: 10.2196/26145] [Medline: 33533727]

Abbreviations

AARC: Awareness of Age-Related Change
CFA: confirmatory factor analysis
CFI: comparative fit index
GFI: goodness-of-fit index
IFI: incremental fit index
MANCOVA: multivariate analysis of covariance
PGFI: parsimony goodness-of-fit index
RMSEA: root mean square error of approximation
TLI: Tucker-Lewis index
TPB: Theory of Planned Behavior
TRQ: Technology Readiness Questionnaire
WHO: World Health Organization

Edited by J Wang; submitted 18.04.24; peer-reviewed by H Shin, M Khamzina; comments to author 18.07.24; revised version received 01.08.24; accepted 14.01.25; published 05.02.25.

Please cite as:
Wong KP, Teh PL, Lim WM, Lee SWH
Enhancing Older Adults' Lives Through Positive Aging Perception, Quality-of-Life Enhancement, and Social Support to Drive Acceptance and Readiness Toward Indoor Assistive Technology: Cross-Sectional Study
JMIR Aging 2025;8:e59665
URL: https://aging.jmir.org/2025/1/e59665
doi:10.2196/59665
PMID:

Original Paper

A Self-Adaptive Serious Game to Improve Motor Learning Among Older Adults in Immersive Virtual Reality: Short-Term Longitudinal Pre-Post Study on Retention and Transfer

Gauthier Everard[1,2,3], PT, MSc, PhD; Louise Declerck[3], PT, MSc, PhD; Thierry Lejeune[3,4,5], MD, PhD; Martin Gareth Edwards[5,6], PsyD, PhD; Justine Bogacki[3], PT; Cléo Reiprich[3], PT; Kelly Delvigne[3], PT; Nicolas Legrain[3], PT; Charles Sebiyo Batcho[1,2], PT, MSc, PhD

[1]School of Rehabilitation Sciences, Faculty of Medicine, Laval University, Québec, QC, Canada

[2]Centre interdisciplinaire de recherche en réadaptation et intégration sociale, Université Laval, Québec, QC, Canada

[3]Neuro Musculo Skeletal Lab (NMSK), Institut de Recherche Expérimentale et Clinique, UCLouvain, Brussels, Belgium

[4]Services de Médecine Physique et Réadaptation, Cliniques Universitaires Saint-Luc, Brussels, Belgium

[5]Louvain Bionics, UCLouvain, Louvain-la-neuve, Belgium

[6]Psychological Sciences Research Institute (IPSY), UCLouvain, Louvain-la-neuve, Belgium

Corresponding Author:
Gauthier Everard, PT, MSc, PhD
School of Rehabilitation Sciences
Faculty of Medicine
Laval University
2325 Rue de l'Université
Québec, QC, G1V0A6
Canada
Phone: 1 4185299141 ext 46914
Email: gauthier.everard@uclouvain.be

Abstract

Background: Despite their potential, the use of serious games within immersive virtual reality (iVR) for enhancing motor skills in older adults remains relatively unexplored. In this study, we developed a self-adaptive serious game in iVR called REAsmash-iVR. This game involves swiftly locating and striking a digital mole presented with various distractors.

Objective: This short-term longitudinal pre-post study aims to evaluate REAsmash-iVR's efficacy in promoting motor learning in older adults. Specifically, we seek to determine the transfer and retention of motor learning achieved through REAsmash-iVR to other iVR tasks.

Methods: A total of 20 older adults participated in the study, engaging with REAsmash-iVR over 7 consecutive days. The evaluation included iVR tests such as KinematicsVR and a VR adaptation of the Box and Block Test (BBT-VR). KinematicsVR tasks included drawing straight lines and circles as fast and as accurately as possible, while BBT-VR required participants to move digital cubes as quickly as possible within 60 seconds. Assessments were conducted before and after the intervention, with a follow-up at 1 week post intervention. The primary outcome focused on evaluating the impact of REAsmash-iVR on speed-accuracy trade-off during KinematicsVR tasks. Secondary outcomes included analyzing movement smoothness, measured by spectral arc length, and BBT-VR scores.

Results: Results revealed significant improvements in speed-accuracy trade-off post intervention compared to that before the intervention, with notable retention of skills for straight lines (t_{19}=5.46; P<.001; Cohen d=1.13) and circle drawing (t_{19}=3.84; P=.001; Cohen d=0.787). Likewise, there was a significant enhancement in spectral arc length, particularly for circle drawing (χ^2_2=11.2; P=.004; ε^2=0.23), but not for straight-line drawing (χ^2_2=2.1; P=.35; ε^2=0.003). Additionally, participants demonstrated transfer with significant improvement (q=5.26; P<.001; Cohen r=0.678) and retention (q=6.82; P<.001; Cohen r=0.880) in BBT-VR skills.

Conclusions: These findings provide perspectives for the use of iVR to improve motor learning in older adults through delivering self-adaptive serious games targeting motor and cognitive functions.

Trial Registration: ClinicalTrials.gov NCT04694833; https://clinicaltrials.gov/study/NCT04694833

(JMIR Aging 2025;8:e64004) doi:10.2196/64004

KEYWORDS

virtual reality; aged; learning; upper extremity; video games; kinematics

Introduction

By 2050, the proportion of older adults (older than 65 years) worldwide will nearly reach 22% [1,2]. Within the population, some older adults experience progressive functional decline, encompassing both motor and cognitive aspects [3]. This decline exacerbates issues of inactivity and sedentariness [4] contributing to increased prevalence of age-related diseases [5,6]. As an illustration, it is projected that by the year 2050, the number of individuals affected by major neurocognitive disorders will surge from 50 to 152 million worldwide, marking a 3-fold increase in cases [7].

Older adults' functional decline is typically associated with lower motor functions [8] and reduced quality of life [9,10]. Consequently, older adults tend to adopt compensatory behaviors that mitigate the impact of these reduced functions on their daily living activities [11]. These behavioral compensations include making slower, less accurate, less linear, and less smooth movements [12,13].

Motor learning refers to any experience-dependent improvement of a skill and typically involves both motor and cognitive processes [14]. Once learned, motor skills can be retained for an extended duration, resulting in sustained enhancements in performance [14]. A skill is not deemed fully acquired until the ability to retain or apply it in different contexts is demonstrated [15]. Research suggests that both healthy older adults and those with neurocognitive disorders can improve their motor performance through motor learning, exhibiting enhancements in movement speed, smoothness, coordination, and accuracy [16,17]. However, current motor learning programs mainly focus on gait and balance and therefore, demand time and availability from caregivers [18]. In response to these demands, new portable devices with cost-saving potential such as virtual reality (VR) might be of interest to promote motor learning in older adults.

VR can be defined as a computerized technological system that allows users to interact with a simulated multisensorial environment while providing real-time performance feedback [19]. Two main types of VR experience exist. Nonimmersive VR (niVR) is where users maintain awareness of their physical surroundings and receive visual feedback via a 2D display. Immersive VR (iVR) facilitates total immersion in the digital environment (using a head-mounted display or a large, curved screen with a panoramic view), with a comprehensive panoramic perspective [20]. Recent research suggests that VR programs may enhance participants' level of physical activity [21] and cognitive skills such as reaction time [22].

In rehabilitation, VR devices are often combined with serious games. Serious games refer to any game-based initiative that primarily focuses on learning objectives (such as education or rehabilitation) rather than simple entertainment [23]. Serious games have the capability to fulfill motor learning principles [24-26], as they motivate participants to make numerous practice repetitions through the use of multisensory feedback, personalized challenges, and through use of compelling and enriched environments [27,28]. After a certain period of familiarization, VR devices may allow participants to follow self-directed interventions and complete remote assessment of objective motor and cognitive performance (eg, analysis of kinematics and reaction time) during interventions [29]. However, despite the potential, the use of serious games in VR to promote motor learning in older adults remains underexplored [30]. In addition, the generalization of skills acquired in iVR to other skills in iVR remains debated.

A recent review has proposed intriguing methods to enhance the comprehension of motor learning in VR, including tracking participants' kinematics, manipulating sensorial feedback and difficulty parameters, and precisely simulating VR physics [27]. Prior work showed that kinematic indexes acquired in iVR (eg, movement linearity) were reliable and could possibly differentiate hand movements between healthy older adults and those with major neurocognitive disorders [31]. Additionally, several studies have provided evidence to support the idea that the provisioning of haptic and visual feedback in VR positively influenced movement smoothness, accuracy, and rapidity, thereby contributing to the improvement of motor learning in VR [32-35]. Recent evidence also highlights the efficacy of incorporating bimanual tasks in VR for promoting unimanual motor learning, aligning with the notion that the acquisition of motor skills in a digital environment can be optimized when the tasks closely mimic the complexities of everyday activities [36-38]. In line with the Yerkes-Dodson law, there is also evidence indicating that to maintain participants' motivation, the level of difficulty should be optimally balanced, neither too hard nor too easy [39]. To achieve this optimal balance, research suggests that game difficulty should be adjustable and tailored to individual participants' motor and cognitive performance [29,40,41]. Research supports the idea that self-adaptive training, where the difficulty is adjusted based on real-time performance, can optimize learning [42]. For instance, 1 study has highlighted that individualized VR training for driving led to more effective learning and retention of performance compared to traditional VR, video, and manual training [43]. The literature indicates success rates ranging from 60% to 80% as ideal for effectively enhancing participants' motivation [40,44].

Regarding older adults, several studies have demonstrated that motor learning in VR can be effective, resulting in improvements in motor performance along with the retention of learned skills over time. For example, a longitudinal study found that both healthy older adults and those with Parkinson disease were able to achieve learning and retention of skills

across 10 different niVR games, with effective transfer of these skills to similar untrained tasks [45]. While these results are promising for niVR, the retention of iVR skills and transfer to other tasks remains underexplored. A recent multicentric large parallel randomized controlled trial (n=293) has nevertheless produced encouraging results, demonstrating that progressive cognitive-motor training in iVR was effective in older adults, with greater improvements in global cognition and physical frailty compared to traditional interventions [46].

Following current recommendations aiming at improving motor learning in older adults, we developed a self-adaptive serious game in iVR (REAsmash-iVR) [47]. This game consists of finding and hitting a digital mole as fast as possible when presented with different types of distractors. In this version, we use a regulator to continuously adapt exercise difficulty according to participants' performance. As this version has not yet been tested among older adults, this work aims to test the feasibility and effectiveness of REAsmash-iVR in promoting motor learning within this population. Specifically, we sought to determine the transfer and retention of motor learning achieved through REAsmash-iVR to other iVR tasks. We hypothesized that REAsmash-iVR would significantly improve participants' unimanual reaching velocity and accuracy in other iVR tasks. We also aimed to assess the effect of REAsmash-iVR on participants' simple reaction time.

Methods

Study Design

This study used a short-term longitudinal pre-post design, with data collected at 3 time points: baseline (T0), immediate postintervention (T1), and 1-week follow-up (T2). This design allowed us to assess both immediate and retained motor learning effects, aligning with established motor learning literature for intermediate-term retention and transfer [39].

Ethical Considerations

The study was conducted in adherence to the principles of the Helsinki Declaration and received approval from the Hospital-Faculty Ethics Committee of Saint-Luc-UCLouvain in Belgium (B403201524184) and the Recherche Sectorielle en Réadaptation et Intégration Sociale Ethics Committee in Canada (#2020-1909). Prior to commencing the trial, all participants provided written informed consent. The study adhered to Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) guidelines. Participation was voluntary and uncompensated, and all data was collected, stored, and analyzed in a manner that ensured participant anonymity.

Participants

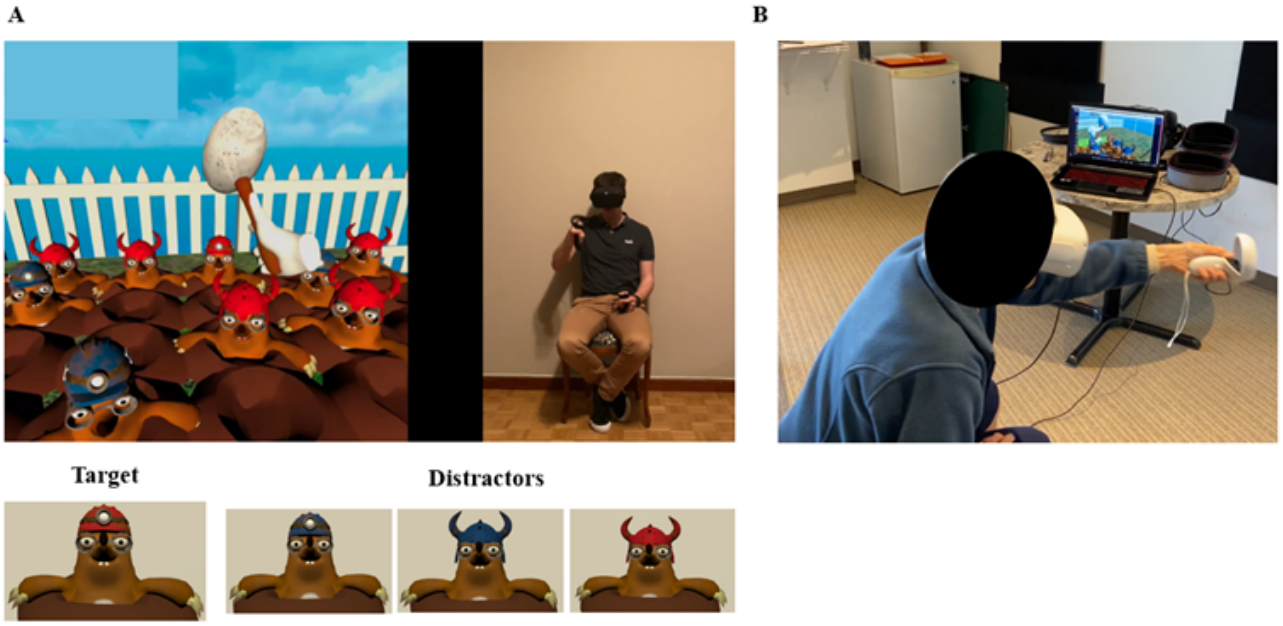
The study recruited participants from the Belgian and Canadian populations between October 2022 and December 2023 using convenient sampling, leveraging word of mouth, and community outreach strategies. Inclusion criteria were individuals aged 65 years and older, possessing corrected-to-normal vision, and demonstrating the ability to comprehend simple instructions. Older adults with orthopedic or neurological disorders that might have impacted their capacity to handle a controller or that could alter upper extremity movements were excluded from the study. Participants' cognition was screened using the Montreal Cognitive Assessment [48].

A flowchart diagram illustrating the participant flow through each stage of the study is presented in the Results section, as recommended by reporting guidelines such as TREND and CONSORT (Consolidated Standards of Reporting Trials) to ensure transparency and clarity in reporting [49,50].

Materials

The self-adaptive serious game REAsmash-iVR was developed using Unity 2019.3.15 software on the Oculus Quest 2 (Meta). This headset provides a high-resolution display (1832×1920 pixels per eye) and up to 90 Hz refresh rate, which ensures smooth and immersive interaction. The system includes 6 degrees of freedom tracking through integrated sensors, which allows participants to move freely within the digital environment. The device is equipped with 2 handheld controllers with motion tracking, which participants use to interact with the game, particularly for striking the target with a digital hammer. The controllers are equipped with motion sensors and buttons for precise input. As presented in Figure 1, Sidequest software was used to facilitate synchronization and video sharing from the headset to a laptop during the experiment.

Figure 1. Illustration of the REAsmash-iVR. (A) The upper part of the panel simultaneously depicts the REAsmash-iVR environment as seen through the VR headset (left) and the corresponding movements performed by a participant while interacting with the system (right). The lower part of the panel shows the target (a mole wearing a red miner helmet) and the distractors (moles wearing blue miner helmets, blue-horned helmets, and red-horned helmets). (B) Experimental setup: This panel illustrates a participant playing REAsmash-iVR using an immersive VR headset connected to a computer, which streams the application. iVR: immersive virtual reality; VR: virtual reality.

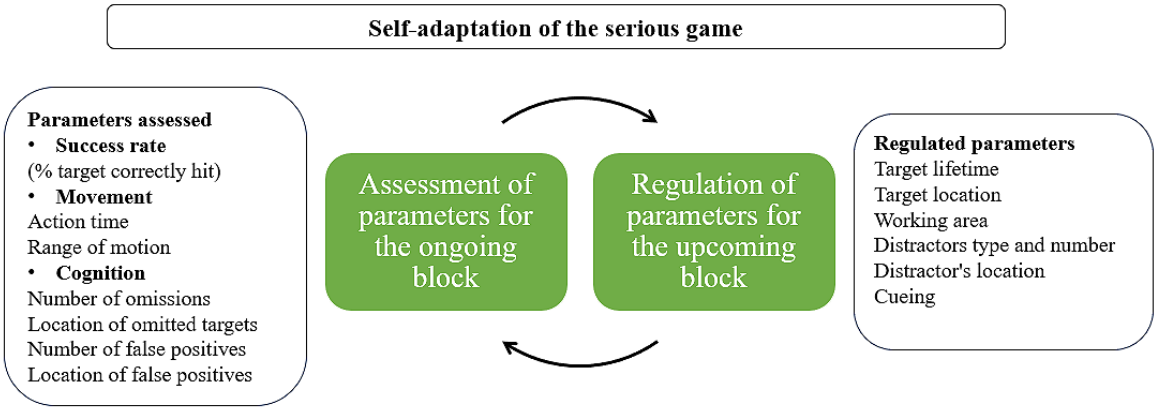


In the REAsmash-iVR serious game, the participant is asked to locate a target as quickly as possible. The target consists of a mole wearing a red miner's helmet and is presented among distractors (moles wearing helmets of different shapes and colors: a blue miner's helmet, a horned blue helmet, and a horned red helmet). These distractors manipulate cognitive difficulty by competing for attention during the task, requiring participants to focus and filter out irrelevant stimuli while searching for the target. To enhance cognitive engagement, distractors were designed to vary in salience, with some mimicking the target more closely in shape and color. Throughout the game, participants were instructed to only hit the target mole with the red miner's helmet, a task that requires both attention and precise motor action. To this end, participants used digital hammers, operated by the game controllers [47]. Motor function difficulty was manipulated through iterative practice of upper extremity reaching motions in different directions and with different levels of velocity.

In the version used here, the REAsmash-iVR used a regulator to adjust difficulty automatically and progressively based on the participants' motor and cognitive performance. The regulator of REAsmash-iVR difficulty aims to have the participant an average 75% successful performance. As the user improved and learned, the game progressively became more difficult, thereby maintaining the 75% optimal success rate. From a motor learning perspective, as the game escalated in difficulty, participants were compelled to execute a greater number of reaching movements toward the moles, spanning further distances, and within shorter timeframes, necessitating more efficient and precise upper extremity actions.

To ensure a continuous and progressive adaptation of the game difficulty, we used a dynamic regulation, with an infinite number of trial blocks. Each block involved finding and hitting a total of 1 to 24 target mole (trials), depending on the participant's level of performance. Between each block, the algorithm moderates the difficulty of the game based on the overall success rate (the ratio between the number of target moles accurately hit and the number of trials) of the prior block. If the success rate was superior to 75%, the game was considered too easy by the algorithm. If the success rate ranged between 50% and 75%, the game was considered difficult. If the success rate was below 50%, the game was considered excessively difficult. Depending on the success rate of the prior block (>75% vs 50%-75% vs <50%), the parameters that were considered responsible for the observed success rate were adjusted by the algorithm. More specifically, as presented in Figure 2, during each block, in addition to the overall success rate, the algorithm evaluated the proportion of omissions (instances when the target mole is not hit vs hit), the location of omitted moles, the number of false positives (distractor moles that were hit), and the location of distractor moles hit. These outcomes were used as indicators by the algorithm to see which of the following parameters had to be adjusted for the next block: the timing and location of the target appearance, the quantity (number) and types of distractors (high vs low salience contrast), the working area (where target and distractors appeared) and the delivery of cues that helped the participant to find the target (no cues, spatial auditive cue, and visual cues).

Figure 2. Self-adaptation of REAsmash-iVR. The system assesses performance parameters (success rate, movement, and cognition) after each block and regulates the task difficulty for the subsequent block by adjusting target and distractor characteristics, working area, and cueing to maintain an optimal challenge level for the participant. iVR: immersive virtual reality.



Experimental Protocol

Participants were tasked with engaging in REAsmash-iVR for 7 consecutive days for at least 15 minutes per day at their home. They were required to maintain a seated position throughout the gameplay and were asked to use both hands during the game. The feasibility of using REAsmash-iVR in older adults encompassed the documentation of adverse events, intervention duration (measured by the total minutes of active interaction with REAsmash-iVR, excluding pauses and time spent on menus and settings), and the evaluation of the game regulation efficacy.

To investigate whether participants' motor learning in REAsmash-iVR transferred to other iVR tasks and demonstrate retention of this transfer, the following measures were performed before the intervention (T0), immediately after (T1), and 7 days later (T2). We assessed motor learning transfer using 2 distinct tools: the KinematicsVR and the iVR version of the Box and Block Test (BBT-VR). These tests were administered by assessors with a background in physiotherapy and experience in using VR. Motor learning transfer refers to the ability to apply skills and improvements acquired during training to new, untrained tasks or contexts. Effective transfer indicates that participants have not only improved performance in the trained task (eg, REAsmash-iVR) but also developed adaptable motor strategies that can be used in other scenarios (eg, KinematicsVR and BBT-VR).

The 1-week duration was chosen as it is commonly used in motor learning literature to assess intermediate-term retention and transfer [39,51]. This timeframe is long enough to observe whether improvements persist beyond the immediate training environment but avoids confounding factors associated with longer intervals, such as unrelated learning or natural recovery. Prior studies have similarly used 1-week retention tests to evaluate skill stabilization and generalization in both laboratory and applied contexts [39,51].

We used the KinematicsVR to assess motor learning transfer in terms of unilateral reaching performance in a novel visuo-motor task [31]. The KinematicsVR requires using one of the VR headset controllers to swiftly and precisely draw 3D shapes (straight lines and a circle) visually presented in an iVR environment. The 3D positions of the controller were registered as an export file (.csv) in the hardware of the headset and analyzed offline. This test was deemed reliable and usable to assess upper extremity kinematics (especially during the drawing of straight lines and circles) in older adults with and without major neurocognitive disorder [31]. For this protocol, participants were asked to perform the movements with their dominant hand. All participants underwent familiarization trials before the assessment. To ensure that the differences between T0 and T1 primarily resulted from the REAsmash-iVR intervention rather than a general learning effect across trials, we contrasted the changes in our sample with those observed in a previous study [31] where an equal sample of older adults underwent the test twice consecutively (with no intervention in between). Transfer in this context indicates that the REAsmash-iVR intervention contributed to general improvements in unilateral reaching performance in a novel task (KinematicsVR). The performance in Kinematics-VR was analyzed based on metrics such as movement velocity, accuracy, and smoothness, allowing us to determine the extent to which the trained skills carried over to this analytic task.

We also used the BBT-VR to evaluate motor learning transfer in terms of gross unilateral manual dexterity [52,53]. This test involved moving digital cubes one at a time from 1 side of a box to the other within a 60-second timeframe. During the test, participants were required to grasp the cubes using their thumb, index, and middle fingers while pressing corresponding buttons on the controller. The BBT-VR was found to be valid, reliable, and usable to assess manual dexterity in healthy adults and individuals with stroke [52,53]. Transfer in this context indicates that the REAsmash-iVR intervention contributed to general improvements in fine motor control and hand-eye coordination. The performance in BBT-VR was analyzed based on metrics such as the number of blocks transferred, allowing us to determine the extent to which the trained skills carried over to this dexterity-focused task.

We also evaluated the participants' motor learning transfer to simple reaction time. The task involved detecting, as quickly as possible, a stimulus presented on a computer screen. Participants were instructed to click as quickly as possible on the touchpad when the stimuli were presented [54].

Kinematic Analyses

By analyzing the kinematic features of movements in KinematicsVR, we could assess whether the movement strategies developed during REAsmash-iVR translated into improved performance in a new visuomotor task.

The 3D positions of the controller obtained during the Kinematics-VR test were extracted from export files (.csv) at a sampling rate of 60 Hz. The analysis of kinematic data was then conducted using a program internally developed in Python (Python Software Foundation). For each participant, a preliminary visual analysis of the data was performed (to ensure that the data were correctly acquired) and signal smoothing was applied using a Butterworth filter (sampling frequency=60 Hz; cutoff frequency=10 Hz). The following kinematic indexes were calculated: the speed-accuracy trade-off (SAT) and the spectral arc length (SPARC).

The SAT, measured in arbitrary units, is a fundamental motor learning metric that quantifies the balance between movement speed and precision, often reflecting the extent to which participants prioritize speed over accuracy or vice versa during task execution. It is widely used to assess training-induced improvements in motor performance, as optimized performance behavior tends to achieve a more favorable balance between these competing demands [55]. In this study, we computed SAT as a ratio between speed and error, using the following equation.

Velocity refers to the first derivative of controller position. Error is measured based on movement linearity, which involves comparing the displacement of the controller with the ideal path. A higher SAT thus reflects a more efficient balance of speed and accuracy, indicating potential motor learning gains.

SPARC is a measure of movement smoothness, which is a key aspect of movement quality and skillful performance. Smoothness, as computed with SPARC, provides critical insights into whether a movement is natural and healthy or involves compensatory strategies. Natural and healthy movements are generally smoother, reflecting efficient neuromotor control, while compensatory movements tend to be less smooth and more erratic. SPARC is computed as the arc length of the instantaneous speed spectrum (ie, the length of the curve depicting the normalized amplitude of the "speed" signal  as a function of its frequency [ω]) [56]. A smoother movement involves less intermittency (alternance of acceleration and deceleration) typically resulting in a more compact and less erratic speed spectrum, leading to a small arc length. A negative sign is added to the computed arc length such a more negative SPARC value corresponds to a less smooth movement. This convention ensures that higher (less negative) SPARC values indicate smoother and more skillful motor performance.

Data Analysis

We performed statistical analyses using Sigmaplot (version; 13.0, Systat Software Inc) with α=.05. For each analysis, we explicitly tested the normality of the data using the Shapiro-Wilk test, and the results informed the selection of appropriate statistical methods (parametric or nonparametric). As this study is the first to test the feasibility and effectiveness of REAsmash-iVR in promoting motor learning, a convenience sample of 20 participants was determined.

To evaluate the efficacy of REAsmash-iVR self-regulation, we reported the percentage of instances where participants achieved a median success rate falling within the range of 60%-80% (a range deemed acceptable for enhancing motivation in a gamified learning context [40,44]). This evaluation was performed across the initial 55 blocks (which represented the minimum number of blocks observed in all participants).

Our primary outcome was to assess participants' motor learning transfer (and retention) to unilateral reaching performance in iVR. To analyze this, we used separate 1-way repeated measure ANOVA (or Friedman test for nonnormal data) for each shape used in the KinematicsVR assessment, comparing participants' SAT across 3 time points: before the intervention (T0), immediately after (T1), and at follow-up (T2). Post hoc pairwise comparisons were performed to detect changes between T0, T1, and T2. We used Bonferroni or Tukey adjustments (depending on the normality of the data) to control for the increased risk of type I error due to multiple comparisons. To ensure that the observed changes in KinematicsVR metrics between T0 and T1 were predominantly attributable to the REAsmash-iVR intervention and not merely a general learning effect over trials, we compared these changes with data from a previous study [31] where older adult participants underwent the test twice consecutively without any intervention in between [31]. To compare the T1-T0 changes between this study and our prior one, we used either the Mann-Whitney rank sum test (for nonnormal data) or the 2-tailed *t* test (for normally distributed data), depending on the distribution of the data.

Secondary outcomes included the assessment of REAsmash-iVR motor learning transfer to movement smoothness, manual dexterity performance in iVR, and simple reaction time. These outcomes were analyzed using 1-way repeated measures ANOVAs (or Friedman tests), with post-hoc pairwise comparisons. We also used Bonferroni or Tukey adjustments, depending on the normality of the data.

The effect size was computed using η^2 for ANOVAs, ε^2 for the Friedman test, and adjusted Cohen *d* for parametric post-hoc pairwise comparisons and 2-tailed *t* tests. For nonparametric tests, Cohen *r* was used. η^2 for ANOVAs and ε^2 for the Friedman test provide an estimate of the effect's magnitude relative to the total variance, with values of 0.01, 0.06, and 0.14 indicating small, medium, and large effects, respectively. Cohen *d* was used to quantify the standardized mean difference between conditions, with values of 0.2, 0.5, and 0.8 representing small, medium, and large effect sizes. For pairwise comparisons following the Friedman test and between-group comparisons with the Mann-Whitney *U* test, Cohen *r* was interpreted as small ($r\approx0.1$), medium ($r\approx0.3$), and large ($r\approx0.5$).

Results

Overview

A total of 20 older adults (of which 9 were women) with a mean age of 77.4 (SD 6.51) years participated in the study. Most of them were right-handed (n=17; 85%). A flowchart diagram is presented in Figure 3. Complementary information on participants' characteristics is provided in Table 1.

Figure 3. The flowchart diagram outlines the study process, including the assessment of 36 older adults for eligibility. Of these, 16 were excluded (7 did not meet the criteria, 7 declined, and 2 for other reasons). Twenty participants received the intervention, completed it, and were analyzed for feasibility, motor learning transfer, and other outcomes. Data were also compared with 20 older adult participants of a previous study.

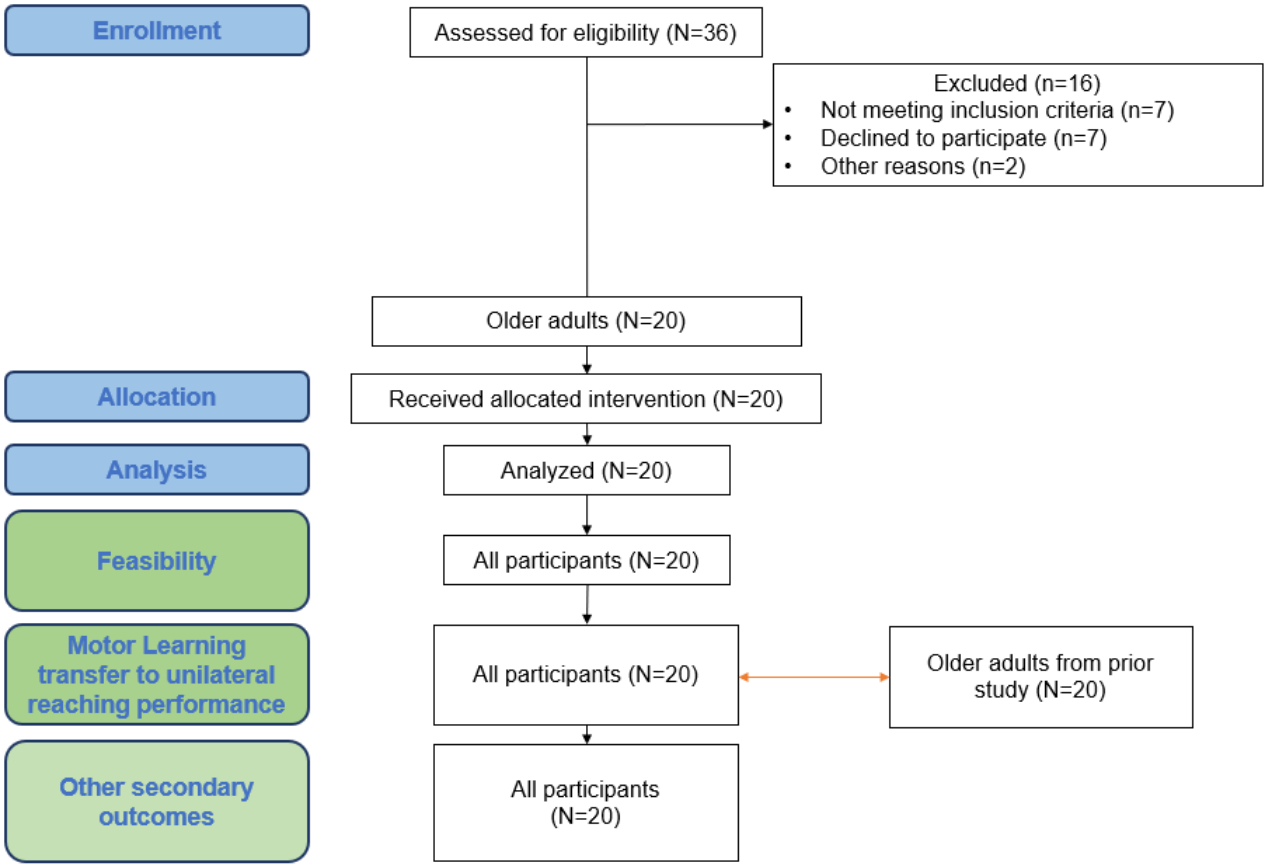


Table 1. Participants' demographics.

Characteristic	Value
Age (years), median (IQR)	77.4 (69 to 89)
Sex, n (%)	
Female	9 (45)
Male	11 (55)
Dominant hand, n (%)	
Right	17 (85)
Left	3 (15)
Height (cm), mean (SD)	169.4 (7.73)
Weight (kg), mean (SD)	77.3 (14.40)
MoCA[a], median (IQR)	25 (6 to 30)
KinematicsVR: Straight lines	
Baseline SAT[b], mean (SD)	2.6 (1.09)
Baseline SPARC[c], median (IQR)	–1.97 (–2.027 to –1.896)
KinematicsVR: Circles	
Baseline SAT, mean (SD)	3.2 (1.51)
Baseline SPARC, mean (SD)	–4.45 (2.227)
Baseline BBT-VR[d], mean (SD)	29 (13.2)

[a]MoCA: Montreal Cognitive Assessment.

[b]SAT: speed-accuracy trade-off.

[c]SPARC: spectral arc length.

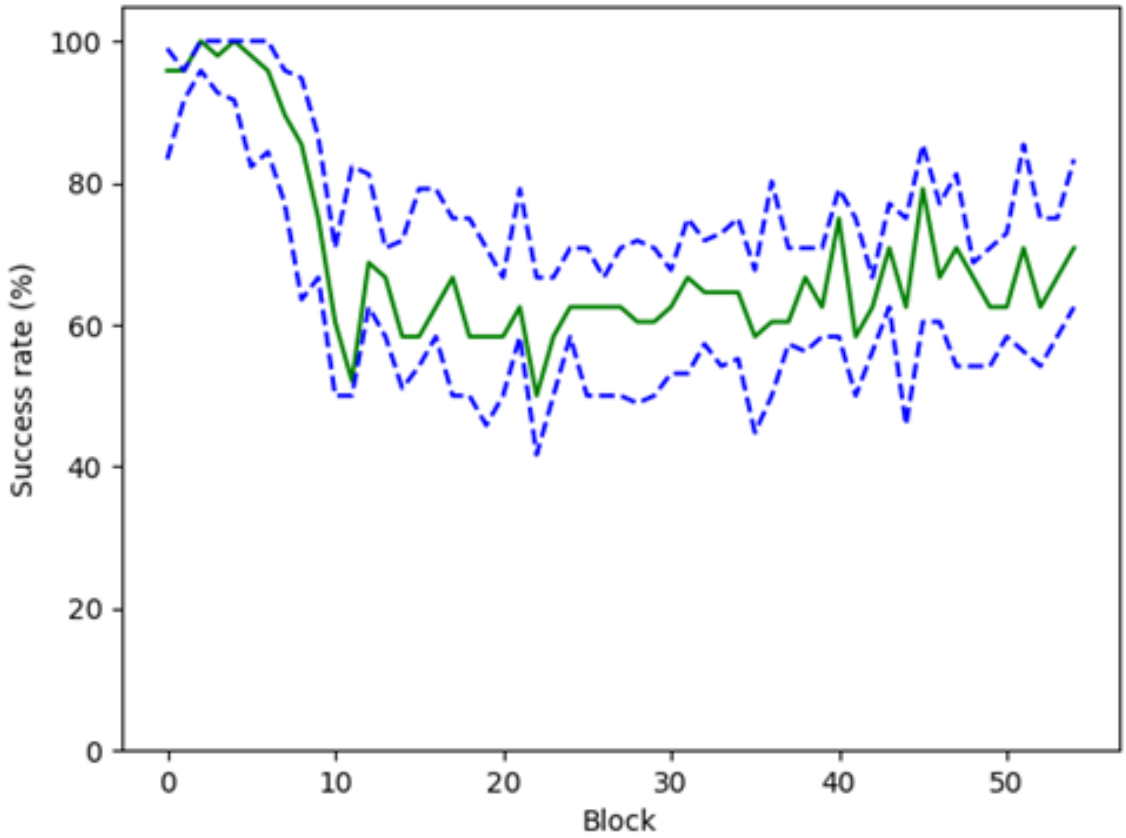
[d]BBT-VR: immersive virtual reality version of the Box and Block Test.

Feasibility

All participants finalized the study, and no adverse event occurred during the intervention. Participants actively played with REAsmash-iVR for 7 consecutive days for a median duration of 15.8 (IQR 9.73-15.12) minutes per day.

As illustrated in Figure 4, participants' median success rate reached a satisfactory value (60%-80%) after completing 11 blocks. Between the 11th and the 55th blocks, participants maintained a median success rate within the 60 to 80% range for 74% of the time (33 out of 44 blocks). Notably, during the final 11 blocks, participants consistently upheld a median success rate between 60% to 80%, reaching 100% coverage.

Figure 4. Evolution of success rate over the blocks. The x-axis represents the block number of the REAsmash-iVR intervention. The y-axis represents the participants' motor success rate. The green line represents the median and the blue lines the 1st and 3rd quartiles issued from all participants. iVR: immersive virtual reality; VR: virtual reality.



Primary Outcome: REAsmash-iVR Motor Learning Transfer to iVR Unilateral Speed-Accuracy Performance

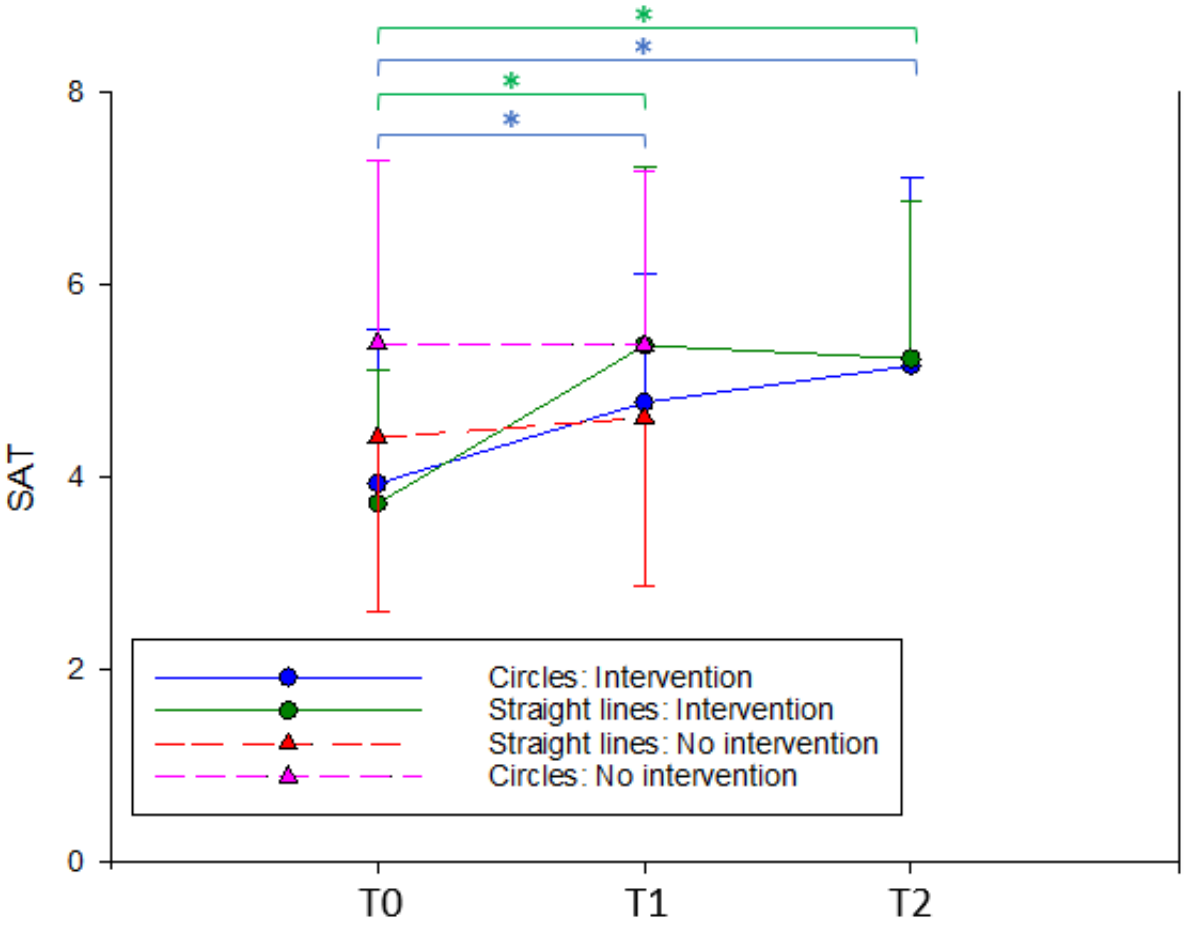
As presented in Table 2 and Figure 5, separate repeated measures ANOVA ($F_{2,38}$=21.9; P<.001; η^2=0.535) and pairwise comparison revealed that directly after the REAsmash-iVR intervention, participants significantly improved their SAT in drawing straight lines (t_{19}=5.97; P<.001; Cohen d=1.20) with retention at T2 (T0 vs T2: t_{19}=5.46; P<.001; Cohen d=1.13). Regarding the drawing of circles, participants showed significant SAT improvements ($F_{2,38}$=7.7; P=.002; η^2=0.290) between T0 and T1 (T0 vs T1: t_{19}=2.64; P=.036; Cohen d=0.613) with retention at T2 (T0 vs T2: t_{19}=3.84; P=.001; Cohen d=0.787; Table 2 and Figure 5).

Table 2. Speed-accuracy trade-off changes over time.

Separate ANOVAs	T0	T1	T2	*F* test (*df*)	RM[a] ANOVA *P* value	Post hoc T0 versus T1 *P* value	Post hoc T0 versus T2 *P* value
Straight lines, mean (SD)	3.7 (1.38)	5.4 (1.85)	5.2 (1.63)	21.9 (2,38)	<.001	<.001	<.001
Circles, mean (SD)	3.9 (1.60)	4.8 (1.34)	5.2 (1.95)	7.7 (2,38)	.002	.036	.001

[a]RM: repeated measures.

Figure 5. Evolution of SAT over time. The * indicates the statistical significance of the ANOVA pairwise comparisons; Results are presented as mean and SDs. SAT: speed-accuracy trade-off.



When drawing straight lines, participants' SAT improvements between T0 and T1 were found to be significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=3.0; *P*=.005; Cohen *d*=0.487). Contrastingly, when drawing circles, participants' SAT improvements between T0 and T1 were not significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=1.96; *P*=.060; Cohen *d*=0.318).

Secondary Outcomes: REAsmash Motor Learning Transfer to iVR Movement Smoothness

As presented in Table 3, the Friedman test and pairwise comparisons showed that for the drawing of straight lines, participants did not observe significant SPARC changes between T0, T1, and T2 (χ^2_2=2.1; *P*=.35; ε^2=0.003; Table 3).

Table 3. Movement smoothness changes over time.

Separate ANOVAs	T0	T1	T2	Chi-square (*df*)	Friedman test *P* value	Post hoc T0 versus T1 *P* value	Post hoc T0 versus T2 *P* value
Straight lines, median (IQR)	–1.97 (–2.027 to –1.896)	–1.92 (–1.978 to –1.783)	–1.91 (–2.016 to –1.811)	2.1 (2)	.35	N/A[a]	N/A
Circles, median (IQR)	–3.77 (–5.876 to –2.517)	–2.67 (–4.000 to –2.359)	–2.44 (–3.287 to –2.273)	11.2 (2)	.004	.03	.004

[a]N/A: not applicable.

Contrastingly, for the drawing of circles, participants showed significant SPARC improvements (χ^2_2=11.20; *P*=.004; ε^2=0.23) between T0 and T1 (*q*=3.58; *P*=.03; Cohen *r*=0.8), with retention at T2 (T0 vs T2: *q*=4.47; *P*=.004; Cohen *r*=1.0; Table 3). Moreover, improvements between T0 and T1 were significantly greater than those observed in participants who did not follow any intervention between the tests (t_{19}=2.41; *P*=0.010; Cohen *d*=0.381).

Secondary Outcomes: REAsmash Motor Learning Transfer to Manual Dexterity Performance

As presented in Table 4, the Friedman test (χ^2_2=25.9; df=2; *P*<.001; ε^2=0.532) and pairwise comparison revealed that participants significantly improved their overall BBT-VR score

between T0 and T1 (q=5.26; P<.001; Cohen r=0.678) with significant retention at T2 (T0 vs T2: q=6.82; P<.001; Cohen r=0.880).

Table 4. BBT-VR[a] scores and reaction times change over time.

	T0	T1	T2	Statistics	RM[b] ANOVA or Friedman test P value	Post hoc T0 versus T1 P value	Post hoc T0 versus T2 P value
BBT-VR score (blocks), median (IQR)	27 (20.1 to 36.4)	36 (29.8 to 45.1)	38 (29.7 to 50.5)	χ^2_2=25.9	<.001	<.001	<.001
Simple reaction time (ms), mean (SD)	398.1 (116.89)	399.9 (94.12)	379.6 (80.31)	$F_{2,38}$=0.5	.64	N/A[c]	N/A

[a]BBT-VR: immersive virtual reality version of the Box and Block Test.

[b]RM: repeated measures.

[c]N/A: not applicable.

Secondary Outcomes: REAsmash Effect on Simple Reaction Time

Participants did not show significant simple reaction time changes ($F_{2,38}$=0.5; P=.64; η^2=0.024) between T0, T1, and T2 (Table 4).

Discussion

Principal Findings

This short-term longitudinal pre-post study aimed to evaluate the feasibility and effectiveness of using a self-adaptive serious game, REAsmash-iVR, to enhance motor learning and simple reaction time in older adults. Our results suggest that, on average, participants required 10 to 40 blocks to benefit from an acceptable-to-optimal level of success rate when starting from the easiest level of difficulty (block 0). Moreover, outcomes indicated that a 7-day intervention with REAsmash-iVR resulted in improved performance in other iVR tasks, as evidenced by enhanced SAT metrics in drawing straight lines and circles, and increased score in displacing digital blocks postintervention. Notably, significant retention of speed/accuracy improvement was observed at the 1-week follow-up for the drawing of straight lines and circles, and for the displacement of digital blocks. Similarly, secondary analyses revealed that the REAsmash-iVR intervention led to improved movement smoothness in drawing circles but not straight lines. Finally, we did not observe any significant effect of REAsmash-iVR on simple reaction time.

REAsmash-iVR Impact on Motor Learning

Our findings seem to indicate that older adults effectively achieved motor learning. After undergoing the REAsmash-iVR intervention, participants exhibited the ability to apply their skills in various contexts, as evidenced by enhanced speed and accuracy in KinematicsVR and improved performance in BBT-VR. Notably, certain improvements persisted even 1 week after the intervention, suggesting intermediate-term retention of learned skills and their potential transfer to other tasks. These findings may highlight a meaningful step in motor learning within this time frame.

On the one hand, our results [31] may indicate that the notable enhancements in speed-accuracy when drawing straight lines postintervention could be credited to the REAsmash-iVR intervention, rather than being merely a result of general learning or increased familiarity with the device and setup. In fact, in our prior study [31], where no intervention occurred between assessments, we observed no significant changes in SAT in KinematicsVR tasks, further supporting the conclusion that the improvements in this study are specifically attributable to the REAsmash-iVR intervention. The self-adaptive nature of REAsmash-iVR, which allowed for tailored and optimized adjustments to accommodate each participant's unique needs and capacities, likely played a role in optimizing engagement and facilitating skill acquisition, ultimately leading to the observed enhancements in motor performance. Previous research has demonstrated that older adults achieved significant motor learning and skill acquisition when provided with appropriate optimized interventions tailored to their specific needs and abilities. For instance, a study comparing a group who practiced a square-stepping task with enhanced feedback, autonomy-supportive choices, and optimized instructions to a control group practicing without these elements, found that the experimental group exhibited faster movement times during both practice and retention phases [57]. Similarly, feedback from participants in our study overwhelmingly attested to their enjoyment of the game. Many remarked on the engaging nature of the REAsmash-iVR intervention, highlighting its immersive qualities and the satisfaction derived from mastering new skills within the digital environment. Furthermore, studies have emphasized the importance of leveraging technological advancements to develop innovative interventions that cater to the unique challenges and preferences of older adults [58,59]. In a prior study, researchers showed that an adaptive video game training intervention led to generalized positive effects on cognitive control abilities in older adults [59]. Our findings also align with [45] a longitudinal, controlled clinical study investigating motor learning, retention, and transfer in older adults using VR-based training, specifically focusing on individuals with Parkinson disease [45]. Their results demonstrated that older adults with Parkinson disease could effectively learn new motor skills through VR-based interventions. Importantly, the study observed improvements not only in motor performance but also in the retention and transfer of learned skills to real-world tasks.

The greater performance improvement here, relative to our previous study [31] could be attributed to the participants

exhibiting lower baseline performance levels in this study, potentially allowing for a greater margin of improvement. This difference becomes particularly apparent when considering circle drawing tasks, where participants had similar baseline performance levels, and improvements in speed and accuracy did not exceed those seen in our prior study [31]. However, this divergence in outcomes could also stem from the inherent differences between drawing straight lines and circles (straight vs cyclic movements). Notably, it could be hypothesized that due to the similarity in movement nature (discrete reaching movements) between drawing straight lines and hitting digital moles, REAsmash-iVR likely played a significant role in augmenting speed and accuracy outcomes in straight-line drawing tasks.

Disparities in Transfer to Movement Smoothness and Manual Dexterity in iVR

Secondary analyses revealed that participants demonstrated enhanced movement smoothness in drawing circles but not straight lines. In a prior study, researchers have recently indicated that the intermittency of movement, as evidenced by the number of velocity peaks, is influenced by the specific task being performed in older adults [60]. In our study, the lack of significant changes observed in the drawing of straight lines could potentially be attributed to the 2D and discrete nature of these movements within the KinematicsVR application. This limitation may have restricted the scope for enhancement. Conversely, when considering the drawing of circles in KinematicsVR, the continuous and 3D nature of the movements offers potential for improvement in an additional dimension (compared to the drawing of straight lines where only 2 dimensions are considered in the KinematicsVR assessment). These hypotheses are supported by the baseline results, wherein participants demonstrated a median SPARC score of –1.97 (IQR –2.027 to –1.896) for straight lines, compared to a median SPARC score of –3.77 (IQR –5.876 to –2.517) for circles. Due to the tendency for smoother movements to be indicated by SPARC values closer to 0, these results underscore a heightened potential for enhancement in circle drawing tasks.

REAsmash-iVR Effect on Simple Reaction Time

The study did not observe any significant effect of REAsmash-iVR on simple reaction time. Although these findings could have been expected, they may reflect the specific design limitations of the intervention and assessment or the need for longer intervention durations to detect changes in these outcomes. The task in iVR involved locating and responding to digital moles as quickly as possible, suggesting the potential for improvements in simple reaction times. Especially since our sample had a mean reaction time (mean 398.1, SD 116.89 ms) more than the normal in equally aged standards [54]. In comparison, in a prior study [54], researchers observed that the average mean reaction time of individuals aged between 61 and 80 years for the same task was 296.1 (SD 63.9) milliseconds. However, REAsmash-iVR engages spatial attention and distractor inhibition, aspects that cannot be adequately assessed solely through simple reaction time measurements, as the latter primarily evaluates alertness levels. Moreover, it is important to consider that reaction time improvements tend to be modest in older adults due to age-related declines in neurological processing speed [61]. Several studies have documented a decline in reaction times with increasing age, reflecting changes in neurological networks and cognitive processing abilities [62-64]. A recent study assessing 861 participants aged 70-90 years also observed that an increase in intraindividual variability of reaction time, considered as a cognitive marker of neurobiological disturbance, was associated with dementia and mortality [65]. Therefore, longer intervention periods and more comprehensive assessments may be necessary to capture subtle improvements in reaction times among older adults participating in VR-based interventions.

Limitations

We acknowledge the following limitations. First, the study's sample size (n=20) and design (pre-post), while appropriate for initial exploration, pose limitations to the generalizability and robustness of the findings. Although a retrospective comparison was used to evaluate the effect of REAsmash-iVR on motor learning, using a randomized controlled trial design would provide stronger evidence for drawing conclusions regarding the intervention's efficacy.

Second, the length of the intervention period may have been too short to observe significant improvements in certain outcomes (eg, reaction time), particularly among healthy participants. A longer study duration would be beneficial for capturing more substantial changes. Future investigations involving healthy older adults and individuals with major neurocognitive conditions could provide valuable insights into the effects of REAsmash-iVR on a broader range of participants.

Third, for the BBT-VR, retrospective comparisons between participants who received the REAsmash-iVR intervention and those who did not were possible. As a result, it remains challenging to definitively attribute observed improvements solely to the intervention itself, as opposed to potential learning effects or mere familiarization with the VR device and testing procedures. Especially since the minimal detectable change of the BBT-VR in healthy adults is relatively high (14.06 for the dominant hand and 18.23 for the nondominant hand) [52]. Therefore, future studies incorporating appropriate control groups are essential to establish the causal relationship between the REAsmash-iVR intervention and the observed transfer of performance.

Implications

The results of this study carry significant clinical and research implications. Clinically, they underscore the potential of VR technology as a novel and engaging approach to promote motor learning and rehabilitation in older adults. Previous research has demonstrated that VR-based interventions can improve motor function and engagement in rehabilitation through gamified experiences, particularly in older populations with age-related declines [66-69]. Health care practitioners working in rehabilitation settings may consider integrating VR-based interventions into their programs targeting motor impairments and age-related declines in physical function. More importantly, the use of self-adaptive serious games such as REAsmash-iVR may offer participants the opportunity for self-rehabilitation

through tailored interventions. These interventions can be specifically designed to target individual motor and cognitive deficits, providing a personalized approach to rehabilitation. Personalization in VR rehabilitation has been shown to improve outcomes by adapting difficulty levels based on real-time performance data [43]. Furthermore, the diverse range of applications available within the VR headset may enable participants to receive real-time feedback on their performance, allowing for continuous monitoring of progress and identifying areas for improvement. Several studies have highlighted the potential of iVR to assess relevant quantitative outcomes, such as hand kinematics, gaze tracking, and reaction time, in a valid and reliable manner [47,70-72]. While traditional assessments may suffer from ceiling or floor effects, limiting their sensitivity to subtle changes in performance, these quantitative metrics provide precise, objective insights into how participants behave during the task [72,73]. This allows clinicians and researchers to track nuanced motor and cognitive responses, facilitating more tailored rehabilitation strategies. Such feedback mechanisms could enhance motivation and engagement, facilitating more effective rehabilitation outcomes. VR may boost motivation by providing immersive environments that promote goal-oriented tasks, immediate feedback, and a sense of achievement through gamified elements [74]. These features not only increase adherence to rehabilitation programs but also foster a positive emotional response, which is critical for sustaining long-term engagement and improving functional recovery [75,76]. Additionally, with advancements in VR and mixed-reality headset technologies, iVR devices hold promise for promoting social interaction and connectivity among older adults [77,78]. Research has shown that digital environments can foster social engagement, reducing isolation and improving mental well-being [79,80]. Digital conferences, collaborative gaming experiences, and social environments can be facilitated through VR platforms, fostering social engagement and reducing feelings of isolation, which are particularly relevant in the context of aging populations and social distancing measures. Incorporating these social aspects into VR-based interventions could not only enhance the overall user experience but would also contribute to the holistic well-being of older adults.

From a research perspective, the study highlights the importance of exploring optimal parameters and mechanisms underlying VR interventions to maximize their therapeutic benefits, eventually using regulators of difficulty. Larger controlled studies are needed to elucidate the long-term effects, optimal dosage, and generalizability of VR interventions across different populations and settings.

Conclusions

This study provides valuable insights into the feasibility and potential effectiveness of using a self-adaptive serious game to enhance motor learning in older adults. While the intervention demonstrated promising results in improving reaching accuracy and velocity balance, and movement smoothness, future research is warranted to elucidate its broader impact on physical and cognitive function in aging populations.

Acknowledgments

The authors thank Stéphane Grade, Camille Ganci, and Anthony Garcia for their technical support in the development of the game. This work was supported by postdoctoral grants of the Centre Interdisciplinaire de recherche en réadaptation et intégration sociale (Cirris), the Réseau provincial de recherche en adaptation-réadaptation, and Wallonie-Bruxelles International (all awarded to GE). The authors also thank the Région Wallonne, the SPW-Economie-Emploi-Recherche, and the Win2Wal Program (convention 1810108) for their support.

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

GE initiated the project. TL, MGE, and GE worked on the development of the application. CSB, LD, and GE designed the protocol. JB, CR, NL, and KD conducted the experiment. GE performed the statistical analyses. GE wrote the manuscript. GE realized the kinematic analyses. TL, MGE, CSB, and LD substantially contributed to the interpretation of the results. TL, LD, MGE, and CSB were major contributors to manuscript editing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wasay M, Grisold W, Carroll W, Shakir R. World Brain Day 2016: celebrating brain health in an ageing population. Lancet Neurol 2016;15(10):1008. [doi: 10.1016/S1474-4422(16)30171-5] [Medline: 27452915]
2. Ageing and health. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=Between%202015%20and%202050%2C%20the,from%2012%25%20to%2022%25 [accessed 2024-02-13]
3. Colón-Emeric CS, Whitson HE, Pavon J, Hoenig H. Functional decline in older adults. Am Fam Physician 2013;88(6):388-394 [FREE Full text] [Medline: 24134046]

4. Botö S, Buvarp DJ, Hansson P, Sunnerhagen KS, Persson CU. Physical inactivity after stroke: incidence and early predictors based on 190 individuals in a 1-year follow-up of the fall study of Gothenburg. J Rehabil Med 2021;53(9):jrm00224 [FREE Full text] [doi: 10.2340/16501977-2852] [Medline: 34121128]
5. Marioni RE, Proust-Lima C, Amieva H, Brayne C, Matthews FE, Dartigues J, et al. Social activity, cognitive decline and dementia risk: a 20-year prospective cohort study. BMC Public Health 2015;15:1089 [FREE Full text] [doi: 10.1186/s12889-015-2426-6] [Medline: 26499254]
6. The Lancet Regional Health-Europe. Securing the future of Europe's ageing population by 2050. Lancet Reg Health Eur 2023;35:100807 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100807] [Medline: 38115962]
7. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
8. Wang Z, Wang J, Guo J, Dove A, Arfanakis K, Qi X, et al. Association of motor function with cognitive trajectories and structural brain differences: a community-based cohort study. Neurology 2023;101(17):e1718-e1728 [FREE Full text] [doi: 10.1212/WNL.0000000000207745] [Medline: 37657942]
9. Pusswald G, Tropper E, Kryspin-Exner I, Moser D, Klug S, Auff E, et al. Health-related quality of life in patients with subjective cognitive decline and mild cognitive impairment and its relation to activities of daily living. J Alzheimers Dis 2015;47(2):479-486. [doi: 10.3233/JAD-150284] [Medline: 26401569]
10. Palacios N, Gao X, Schwarzschild M, Ascherio A. Declining quality of life in Parkinson disease before and after diagnosis. J Parkinsons Dis 2012;2(2):153-160 [FREE Full text] [doi: 10.3233/JPD-2012-12083] [Medline: 23939439]
11. Tomaszewski Farias S, Schmitter-Edgecombe M, Weakley A, Harvey D, Denny KG, Barba C, et al. Compensation strategies in older adults: association with cognition and everyday function. Am J Alzheimers Dis Other Demen 2018;33(3):184-191 [FREE Full text] [doi: 10.1177/1533317517753361] [Medline: 29357670]
12. Yan JH, Rountree S, Massman P, Doody RS, Li H. Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. J Psychiatr Res 2008;42(14):1203. [doi: 10.1016/j.jpsychires.2008.01.006] [Medline: 18280503]
13. Yan JH. Effects of aging on linear and curvilinear aiming arm movements. Exp Aging Res 2000;26(4):393-407. [doi: 10.1080/036107300750015778] [Medline: 11091944]
14. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor learning. Compr Physiol 2019;9(2):613-663. [doi: 10.1002/cphy.c170043] [Medline: 30873583]
15. Muratori LM, Lamberg EM, Quinn L, Duff SV. Applying principles of motor learning and control to upper extremity rehabilitation. J Hand Ther 2013;26(2):94-103 [FREE Full text] [doi: 10.1016/j.jht.2012.12.007] [Medline: 23598082]
16. Yan JH, Dick MB. Practice effects on motor control in healthy seniors and patients with mild cognitive impairment and alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2006;13(3-4):385-410. [doi: 10.1080/138255890969609] [Medline: 16887780]
17. Onushko T, Kim C, Christou EA. Reducing task difficulty during practice improves motor learning in older adults. Exp Gerontol 2014;57:168-174. [doi: 10.1016/j.exger.2014.06.006] [Medline: 24909352]
18. Shafizadeh M, Farina N, Parvonpour S. Effects of motor learning interventions on walking performance and physical function in older adults with cognitive impairment and dementia: a systematic review and meta-analysis. J Aging Phys Act 2023;31(2):352-363. [doi: 10.1123/japa.2021-0442] [Medline: 35981711]
19. Saposnik G, Levin M, Outcome Research Canada (SORCan) Working Group. Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. Stroke 2011;42(5):1380-1386. [doi: 10.1161/STROKEAHA.110.605451] [Medline: 21474804]
20. Huygelier H, Mattheus E, Abeele VV, van Ee R, Gillebert CR. The use of the term virtual reality in post-stroke rehabilitation: a scoping review and commentary. Psychol Belg 2021;61(1):145-162 [FREE Full text] [doi: 10.5334/pb.1033] [Medline: 34131490]
21. Ng YL, Ma F, Ho FK, Ip P, Fu KW. Effectiveness of virtual and augmented reality-enhanced exercise on physical activity, psychological outcomes, and physical performance: a systematic review and meta-analysis of randomized controlled trials. Comput Hum Behav 2019;99:278-291. [doi: 10.1016/j.chb.2019.05.026]
22. Rutkowski S, Adamczyk M, Pastuła A, Gos E, Luque-Moreno C, Rutkowska A. Training using a commercial immersive virtual reality system on hand-eye coordination and reaction time in young musicians: a pilot study. Int J Environ Res Public Health 2021;18(3):1297 [FREE Full text] [doi: 10.3390/ijerph18031297] [Medline: 33535539]
23. Mubin O, Alnajjar F, Al Mahmud A, Jishtu N, Alsinglawi B. Exploring serious games for stroke rehabilitation: a scoping review. Disabil Rehabil Assist Technol 2022;17(2):159-165. [doi: 10.1080/17483107.2020.1768309] [Medline: 32508187]
24. Doumas I, Everard G, Dehem S, Lejeune T. Serious games for upper limb rehabilitation after stroke: a meta-analysis. J Neuroeng Rehabil 2021;18(1):100 [FREE Full text] [doi: 10.1186/s12984-021-00889-1] [Medline: 34130713]
25. Maier M, Rubio Ballester B, Duff A, Duarte Oller E, Verschure PFMJ. Effect of specific over nonspecific VR-based rehabilitation on poststroke motor recovery: a systematic meta-analysis. Neurorehabil Neural Repair 2019;33(2):112-129 [FREE Full text] [doi: 10.1177/1545968318820169] [Medline: 30700224]

26. Demers M, Fung K, Subramanian SK, Lemay M, Robert MT. Integration of motor learning principles into virtual reality interventions for individuals with cerebral palsy: systematic review. JMIR Serious Games 2021;9(2):e23822 [FREE Full text] [doi: 10.2196/23822] [Medline: 33825690]
27. Levac DE, Huber ME, Sternad D. Learning and transfer of complex motor skills in virtual reality: a perspective review. J Neuroeng Rehabil 2019;16(1):121 [FREE Full text] [doi: 10.1186/s12984-019-0587-8] [Medline: 31627755]
28. Saragih ID, Everard G, Lee B. A systematic review and meta-analysis of randomized controlled trials on the effect of serious games on people with dementia. Ageing Res Rev 2022;82:101740. [doi: 10.1016/j.arr.2022.101740] [Medline: 36191740]
29. Levin MF, Weiss PL, Keshner EA. Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. Phys Ther 2015;95(3):415-425 [FREE Full text] [doi: 10.2522/ptj.20130579] [Medline: 25212522]
30. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020;3(2):e17286 [FREE Full text] [doi: 10.2196/17286] [Medline: 32780020]
31. Everard G, Boivin S, Boulay G, Duchemin R, Batcho CS. Immersive virtual reality to assess arm kinematics among older adults with and without major neurocognitive disorder—an exploratory cross-sectional study. Neuroscience 2024;537:47-57. [doi: 10.1016/j.neuroscience.2023.10.024] [Medline: 38006964]
32. Krogmeier C, Mousas C, Whittinghill D. Human-virtual character interaction: toward understanding the influence of haptic feedback. Comput Animat Virtual 2019;30(3-4):e1883. [doi: 10.1002/cav.1883]
33. Lontschar S, Deegan D, Humer I, Pietroszek K, Eckhardt C. Analysis of haptic feedback and its influences in virtual reality learning environments. 2020 Presented at: 6th International Conference of the Immersive Learning Research Network (iLRN); June 21-25, 2020; San Luis Obispo, CA. [doi: 10.23919/ilrn47897.2020.9155087]
34. Bugdadi A, Sawaya R, Bajunaid K, Olwi D, Winkler-Schwartz A, Ledwos N, et al. Is virtual reality surgical performance influenced by force feedback device utilized? J Surg Educ 2019;76(1):262-273. [doi: 10.1016/j.jsurg.2018.06.012] [Medline: 30072262]
35. Everard GJ, Lejeune TM, Batcho CS. Visual feedback and age affect upper limb reaching accuracy and kinematics in immersive virtual reality among healthy adults. Int J Rehabil Res 2023;46(3):221-229. [doi: 10.1097/MRR.0000000000000588] [Medline: 37334800]
36. Hoare B, Greaves S. Unimanual versus bimanual therapy in children with unilateral cerebral palsy: same, same, but different. J Pediatr Rehabil Med 2017;10(1):47-59. [doi: 10.3233/PRM-170410] [Medline: 28339410]
37. Kantak S, Jax S, Wittenberg G. Bimanual coordination: A missing piece of arm rehabilitation after stroke. Restor Neurol Neurosci 2017;35(4):347-364. [doi: 10.3233/RNN-170737] [Medline: 28697575]
38. Bhakuni R, Mutha PK. Learning of bimanual motor sequences in normal aging. Front Aging Neurosci 2015;7:76 [FREE Full text] [doi: 10.3389/fnagi.2015.00076] [Medline: 26005417]
39. Schmidt RA, Lee TD, Winstein C, Wulf G, Zelaznik HN. Motor Control and Learning: A Behavioral Emphasis. 6th edition. Champaign, IL: Human Kinetics; 2018.
40. Cameirão MS, Badia SBI, Oller ED, Verschure PFMJ. Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. J Neuroeng Rehabil 2010;7:48 [FREE Full text] [doi: 10.1186/1743-0003-7-48] [Medline: 20860808]
41. Heins S, Dehem S, Montedoro V, Dehez B, Edwards M, Stoquart G. Robotic-assisted serious game for motor and cognitive post-stroke rehabilitation. 2017 Presented at: IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH); April 2-4, 2017; Perth, WA, Australia.
42. Zahabi M, Abdul Razak AM. Adaptive virtual reality-based training: a systematic literature review and framework. Virtual Reality 2020;24(4):725-752. [doi: 10.1007/s10055-020-00434-w]
43. Lang Y, Wei L, Xu F, Zhao Y. Synthesizing personalized training programs for improving driving habits via virtual reality. 2018 Presented at: IEEE Conference on Virtual Reality and 3D User Interfaces (VR); March 18-22, 2018; Tuebingen/Reutlingen, Germany.
44. van der Kooij K, In 't Veld L, Hennink T. Motivation as a function of success frequency. Motiv Emot 2021;45(6):759-768 [FREE Full text] [doi: 10.1007/s11031-021-09904-3] [Medline: 34608344]
45. dos Santos Mendes FA, Pompeu JE, Modenesi Lobo A, Guedes da Silva K, Oliveira TDP, Peterson Zomignani A, et al. Motor learning, retention and transfer after virtual-reality–based training in Parkinson's disease—effect of motor and cognitive demands of games: a longitudinal, controlled clinical study. Physiotherapy 2012;98(3):217-223. [doi: 10.1016/j.physio.2012.06.001] [Medline: 22898578]
46. Kwan RYC, Liu J, Sin OSK, Fong KNK, Qin J, Wong JCY, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024;26:e57809 [FREE Full text] [doi: 10.2196/57809] [Medline: 39259959]
47. Ajana K, Everard G, Lejeune T, Edwards MG. A feature and conjunction visual search immersive virtual reality serious game for measuring spatial and distractor inhibition attention using response time and action kinematics. J Clin Exp Neuropsychol 2023;45(3):292-303. [doi: 10.1080/13803395.2023.2218571] [Medline: 37260369]

48. Hobson J. The Montreal Cognitive Assessment (MoCA). Occup Med 2015;65(9):764-765. [doi: 10.1093/occmed/kqv078] [Medline: 26644445]
49. Haynes AB, Haukoos JS, Dimick JB. TREND reporting guidelines for nonrandomized/quasi-experimental study designs. JAMA Surg 2021;156(9):879-880. [doi: 10.1001/jamasurg.2021.0552] [Medline: 33825826]
50. Hopewell S, Boutron I, Chan A, Collins GS, de Beyer JA, Hróbjartsson A, et al. An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials. Nat Med 2022;28(9):1740-1743. [doi: 10.1038/s41591-022-01989-8] [Medline: 36109642]
51. Walker MP, Stickgold R. Sleep-dependent learning and memory consolidation. Neuron 2004;44(1):121-133 [FREE Full text] [doi: 10.1016/j.neuron.2004.08.031] [Medline: 15450165]
52. Everard G, Otmane-Tolba Y, Rosselli Z, Pellissier T, Ajana K, Dehem S, et al. Concurrent validity of an immersive virtual reality version of the Box and Block Test to assess manual dexterity among patients with stroke. J Neuroeng Rehabil 2022;19(1):7 [FREE Full text] [doi: 10.1186/s12984-022-00981-0] [Medline: 35065678]
53. Everard G, Burton Q, Van de Sype V, Bibentyo TN, Auvinet E, Edwards MG, et al. Extended reality to assess post-stroke manual dexterity: contrasts between the classic box and block test, immersive virtual reality with controllers, with hand-tracking, and mixed-reality tests. J Neuroeng Rehabil 2024;21(1):36 [FREE Full text] [doi: 10.1186/s12984-024-01332-x] [Medline: 38491540]
54. Deary IJ, Liewald D, Nissan J. A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. Behav Res Methods 2011;43(1):258-268. [doi: 10.3758/s13428-010-0024-1] [Medline: 21287123]
55. FITTS PM. The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol 1954;47(6):381-391. [Medline: 13174710]
56. Balasubramanian S, Melendez-Calderon A, Roby-Brami A, Burdet E. On the analysis of movement smoothness. J Neuroeng Rehabil 2015;12:112 [FREE Full text] [doi: 10.1186/s12984-015-0090-9] [Medline: 26651329]
57. Khalaji Z, Nezakat Alhosseini M, Safavi Hamami S, Iwatsuki T, Wulf G. Optimizing Motor Learning in Older Adults. J Gerontol B Psychol Sci Soc Sci 2024;79(1):gbad120. [doi: 10.1093/geronb/gbad120] [Medline: 37591686]
58. Fischl C, Blusi M, Lindgren H, Nilsson I. Tailoring to support digital technology-mediated occupational engagement for older adults - a multiple case study. Scand J Occup Ther 2020;27(8):577-590 [FREE Full text] [doi: 10.1080/11038128.2020.1760347] [Medline: 32396419]
59. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486] [Medline: 24005416]
60. Gulde P, Schmidle S, Aumüller A, Hermsdörfer J. The effects of speed of execution on upper-limb kinematics in activities of daily living with respect to age. Exp Brain Res 2019;237(6):1383-1395. [doi: 10.1007/s00221-019-05507-0] [Medline: 30887078]
61. Kempermann G, Gast D, Gage FH. Neuroplasticity in old age: sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. Ann Neurol 2002;52(2):135-143. [doi: 10.1002/ana.10262] [Medline: 12210782]
62. Der G, Deary IJ. Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. Psychol Aging 2006;21(1):62-73. [doi: 10.1037/0882-7974.21.1.62] [Medline: 16594792]
63. Williams RS, Biel AL, Wegier P, Lapp LK, Dyson BJ, Spaniol J. Age differences in the Attention Network Test: Evidence from behavior and event-related potentials. Brain Cogn 2016;102:65-79. [doi: 10.1016/j.bandc.2015.12.007] [Medline: 26760449]
64. Novikov NI, Brazhnik ES, Kitchigina VF. Pathological correlates of cognitive decline in Parkinson's disease: from molecules to neural networks. Biochemistry 2023;88(11):1890-1904. [doi: 10.1134/S0006297923110172] [Medline: 38105206]
65. Kochan NA, Bunce D, Pont S, Crawford JD, Brodaty H, Sachdev PS. Is intraindividual reaction time variability an independent cognitive predictor of mortality in old age? Findings from the Sydney Memory and Ageing Study. PLoS One 2017;12(8):e0181719 [FREE Full text] [doi: 10.1371/journal.pone.0181719] [Medline: 28792946]
66. Zhu S, Sui Y, Shen Y, Zhu Y, Ali N, Guo C, et al. Effects of virtual reality intervention on cognition and motor function in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Front Aging Neurosci 2021;13:586999 [FREE Full text] [doi: 10.3389/fnagi.2021.586999] [Medline: 34025384]
67. Bixter M, Blocker K, Rogers W. Enhancing social engagement of older adults through technology. In: Pak R, McLaughlin AC, editors. Aging, Technology and Health. London, UK: Academic Press; 2018:179-214.
68. Kalantari S, Bill Xu T, Mostafavi A, Lee A, Barankevich R, Boot WR, et al. Using a nature-based virtual reality environment for improving mood states and cognitive engagement in older adults: a mixed-method feasibility study. Innov Aging 2022;6(3):igac015 [FREE Full text] [doi: 10.1093/geroni/igac015] [Medline: 35592668]
69. Arlati S, Di Santo SG, Franchini F, Mondellini M, Filiputti B, Luchi M, et al. Acceptance and usability of immersive virtual reality in older adults with objective and subjective cognitive decline. J Alzheimers Dis 2021;80(3):1025-1038. [doi: 10.3233/JAD-201431] [Medline: 33646164]
70. Llanes-Jurado J, Marín-Morales J, Guixeres J, Alcañiz M. Development and calibration of an eye-tracking fixation identification algorithm for immersive virtual reality. Sensors 2020;20(17):4956 [FREE Full text] [doi: 10.3390/s20174956] [Medline: 32883026]

71. Campo-Prieto P, Cancela-Carral JM, Rodríguez-Fuentes G. Immersive virtual reality reaction time test and relationship with the risk of falling in Parkinson's disease. Sensors 2023;23(9):4529 [FREE Full text] [doi: 10.3390/s23094529] [Medline: 37177733]
72. Negu A, Matu S, Sava FA, David D. Virtual reality measures in neuropsychological assessment: a meta-analytic review. Clin Neuropsychol 2016;30(2):165-184. [doi: 10.1080/13854046.2016.1144793] [Medline: 26923937]
73. Hao J, Li Y, Remis A, He Z, Yao Z, Pu Y. Performance-based outcome measures of upper extremity in virtual reality and telerehabilitation: a systematic review. Neurol Sci 2024;45(3):977-986. [doi: 10.1007/s10072-023-07158-0] [Medline: 37904015]
74. Aderinto N, Olatunji G, Abdulbasit MO, Edun M, Aboderin G, Egbunu E. Exploring the efficacy of virtual reality-based rehabilitation in stroke: a narrative review of current evidence. Ann Med 2023;55(2):2285907 [FREE Full text] [doi: 10.1080/07853890.2023.2285907] [Medline: 38010358]
75. Buyle M, Jung Y, Pavlou M, Gonzalez SC, Bamiou D. The role of motivation factors in exergame interventions for fall prevention in older adults: a systematic review and meta-analysis. Front Neurol 2022;13:903673 [FREE Full text] [doi: 10.3389/fneur.2022.903673] [Medline: 35989930]
76. Resnick B. Efficacy beliefs in geriatric rehabilitation. J Gerontol Nurs 1998;24(7):34-44. [doi: 10.3928/0098-9134-19980701-08] [Medline: 9801529]
77. Hughes S, Warren-Norton K, Spadafora P, Tsotsos L. Supporting optimal aging through the innovative use of virtual reality technology. Multimodal Technol Interact 2017;1(4):23. [doi: 10.3390/mti1040023]
78. Healy D, Flynn A, Conlan O, McSharry J, Walsh J. Older adults' experiences and perceptions of immersive virtual reality: systematic review and thematic synthesis. JMIR Serious Games 2022;10(4):e35802 [FREE Full text] [doi: 10.2196/35802] [Medline: 36472894]
79. Lin CX, Lee C, Lally D, Coughlin JF. Impact of virtual reality (VR) experience on older adults' well-being. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment. Cham, Switzerland: Springer; 2018.
80. Matsangidou M, Solomou T, Frangoudes F, Papayianni E, Pattichis CS. Offering outworld experiences to in-patients with dementia through virtual reality: mixed methods study. JMIR Aging 2023;6:e45799 [FREE Full text] [doi: 10.2196/45799] [Medline: 37656031]

Abbreviations

BBT-VR: immersive virtual reality version of the Box and Block Test
CONSORT: Consolidated Standards of Reporting Trials
iVR: immersive virtual reality
niVR: nonimmersive virtual reality
SAT: speed-accuracy trade-off
SPARC: spectral arc length
TREND: Transparent Reporting of Evaluations with Nonrandomized Designs
VR: virtual reality

Edited by J Wang; submitted 05.07.24; peer-reviewed by S Balula, L Wang, S Pardini; comments to author 11.11.24; revised version received 13.12.24; accepted 30.01.25; published 03.03.25.

Please cite as:
Everard G, Declerck L, Lejeune T, Edwards MG, Bogacki J, Reiprich C, Delvigne K, Legrain N, Batcho CS
A Self-Adaptive Serious Game to Improve Motor Learning Among Older Adults in Immersive Virtual Reality: Short-Term Longitudinal Pre-Post Study on Retention and Transfer
JMIR Aging 2025;8:e64004
URL: https://aging.jmir.org/2025/1/e64004
doi:10.2196/64004
PMID:40053708

Original Paper

Socioculturally Appropriate Internet-Based Geriatric Care Model for Older Adults Living With HIV: Experience-Based Co-Design Approach

Kristina M Kokorelias[1,2,3,4], PhD; Marina B Wasilewski[2,3,5], PhD; Dean Valentine[1]; Andrew D Eaton[6,7], MSW, PhD; Erica Dove[3], MSc; Paige Brown[8], BSc; Stuart McKinlay[8], BHSc; Christine L Sheppard[7], MSW, PhD; Esther Su[1]; Hardeep K Singh[2,3,9], PhD; Ashley Flanagan[4], PhD; Alice Zhabokritsky[10,11,12], MD; Reham Abdelhalim[13], MD, PhD; Rabea Parpia[14], MD; Rahel Zewude[11], MD; Laura Jamieson[15], MSc; Sharon Walmsley[10,11,12], MSc, MD; Luxey Sirisegaram[1], MD

[1]Section of Geriatric Medicine, Department of Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada
[2]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada
[3]Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada
[4]National Institute on Ageing, Toronto Metropolitan University, Toronto, ON, Canada
[5]St John's Rehab Research Program, Sunnybrook Health Sciences Centre, Toronto, ON, Canada
[6]Faculty of Social Work, Saskatoon Campus, University of Regina, Saskatoon, SK, Canada
[7]Factor-Inwentash School of Social Work, University of Toronto, Toronto, ON, Canada
[8]Undergraduate Medical Education, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[9]KITE, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[10]Infectious Diseases, Department of Medicine, University Health Network, Toronto, ON, Canada
[11]Division of Infectious Diseases, Department of Medicine, University of Toronto, Toronto, ON, Canada
[12]CIHR Canadian HIV Trials Network, Vancouver, BC, Canada
[13]Burlington Ontario Health Team, Joseph Brant Memorial Hospital, Burlington, ON, Canada
[14]St Michael's Hospital, Toronto, ON, Canada
[15]Ontario Federation of Indigenous Friendship Centres, Toronto, ON, Canada

Corresponding Author:
Kristina M Kokorelias, PhD
Section of Geriatric Medicine
Department of Medicine
Sinai Health System and University Health Network
Suite 18-208
600 University Avenue
Toronto, ON, M5G 1X5
Canada
Phone: 1 416 586 4800 ext 4374
Email: Kristina.kokorelias@sinaihealth.ca

Abstract

Background: Older adults living with HIV face challenges accessing regular geriatric care, and while virtual care services could offer a solution, they may come with limitations.

Objective: This study aimed to co-design a culturally appropriate virtual care model tailored to older adults' needs using the experience-based co-design methodology.

Methods: We used a qualitative, experience-based co-design approach with 19 older adults living with HIV. The process involved 3 phases: identifying needs through interviews and questionnaires, codeveloping a care model prototype through focus groups and a workshop, and refining the model using feedback from a world café format. Data were analyzed using thematic content analysis.

Results: The co-design process led to a virtual care model prototype that directly addressed participants' key needs. These included personalized communication methods, simplified technology interfaces for easier access, and culturally responsive care practices. Participants emphasized the importance of privacy in virtual consultations, flexible scheduling to accommodate health fluctuations, and ongoing support for managing both HIV and aging-related conditions. Their feedback shaped a model designed to bridge service gaps, offering a more inclusive, accessible, and patient-centered approach to virtual geriatric care.

Conclusions: This study co-designed a potential virtual geriatric care model grounded in the experiences of older adults living with HIV. By integrating participants' insights throughout the design process, the model offers a promising approach to improving care for this vulnerable population. Future directions for research to test this model are proposed.

(JMIR Aging 2025;8:e67122) doi:10.2196/67122

KEYWORDS

older adults; HIV; virtual care; co-design; qualitative research

Introduction

Background

The global population is aging, with the number of people aged ≥50 years projected to double by 2050, reaching approximately 2 billion individuals worldwide [1]. The aging global population also includes a growing number of older adults living with HIV [2]. Due to advancements in combination antiretroviral therapy, the life expectancy of people living with HIV has significantly increased [3], leading to the first generation of people living with HIV now reaching geriatric age [4]. This demographic shift brings unique challenges as older adults living with HIV experience a complex interplay of age-related comorbidities and HIV-related health issues [5-7].

Older adults living with HIV encounter a myriad of challenges that significantly impact their geriatric care. Medically, these individuals are more likely to experience comorbid conditions, including cardiovascular disease [6,8,9], diabetes [6,10], and osteoporosis [11-13], which can complicate HIV management and treatment [14]. These comorbidities complicate treatment regimens and require coordinated care, which may not always be available [15,16]. The specialization of health care often leads to fragmented care, where health care providers focus solely on HIV management without considering the broader geriatric care needs of the patient, such as mobility, nutrition, and mental health [17,18]. Furthermore, the physiological changes associated with aging, combined with the long-term effects of HIV and its treatment, can exacerbate these comorbidities, leading to a higher burden of illness and increased health care use [19]. Psychologically, older adults living with HIV may experience higher rates of depression [20-22], anxiety [23,24], and cognitive impairment [25-27] than their younger counterparts living with HIV. The stigma associated with both aging and HIV can combine and contribute to social isolation, discrimination in health care settings, and declines in mental health and overall well-being [28-30]. In addition, older adults may encounter ageism within health care settings, which can negatively affect the quality of care they receive and their willingness to seek further treatment [31-33]. Thus, the increase in older people living with HIV underscores the increasing need for comprehensive geriatric care that addresses the complex health needs of older adults [34,35].

HIV is highly stigmatized, and this can exacerbate other experiences of stigma and discrimination [32,36,37]. Notably, geriatric care (ie, care delivered specifically for older adults) is most effective when it meets the sociocultural needs of its patients [38,39]. Socioculturally appropriate care refers to health care that respects and responds to the cultural and social needs of patients, integrating their backgrounds, beliefs, and values into their care plans [40-43]. This approach is crucial for improving patient outcomes, enhancing satisfaction with care, and reducing health disparities [44-46]. However, delivering socioculturally appropriate care, particularly in the intersection of geriatrics and HIV, presents unique challenges [43]. These include a lack of cultural competency training [47,48], insufficient resources [49-51], language barriers [52,53], and implicit biases among health care providers [54-56], which can lead to misunderstandings and suboptimal care [57]. Moreover, individuals aging with HIV may come from diverse countries and have varying sexual identities and unique histories of trauma, discrimination, and stigma that differ from those of the general population [36,37,58]. Cultural mismatches such as failing to address the stigma associated with HIV in certain cultures can result in reduced patient engagement, lower treatment adherence, and worse health outcomes [36,59]. Therefore, addressing these challenges is essential to provide effective geriatric care to diverse people living with HIV.

Virtual care holds significant potential to address some of the challenges faced by older people living with HIV [60-63]. For instance, virtual care can help facilitate regular check-ins and monitoring without the need for frequent in-person clinic visits, which is particularly beneficial for those with mobility issues or transportation obstacles or for those living in remote areas [64-66]. As one example, virtual interventions for rural older veterans living with HIV have been shown to improve access to care and patient health outcomes, including quality of life, medication adherence, depressive symptoms, internalized stigma, and loneliness [67]. Virtual platforms can also provide a discrete way for patients to access care, potentially reducing stigma-related barriers to seeking health care in person [68-70].

While virtual care has revolutionized health care delivery, it also presents several challenges that need to be addressed for effective implementation [71,72]. One major issue is the digital divide, where limited access to technology and reliable internet connectivity, especially in rural and underserved areas, hampers the ability of patients to engage in virtual care [73-75]. In

addition, digital literacy poses a significant barrier as many patients, particularly older adults, may lack the knowledge and skills to navigate digital health platforms effectively [17,73,76]. There is a need to develop specialized socioculturally appropriate virtual geriatric care models tailored to the unique needs of diverse older people, including those living with HIV. Geriatric care refers to the specialized medical and supportive care tailored to meet the complex health and social needs of older adults, encompassing chronic disease management, functional support, and holistic well-being [50].

Delivering geriatric care virtually for people living with HIV presents both opportunities and significant challenges. While virtual care can enhance access to health care for older adults living with HIV, particularly in underserved or rural areas, it also requires careful consideration of various factors. One of the primary challenges is the complexity of managing multimorbidity, a hallmark of aging in people living with HIV, which often involves the treatment of multiple chronic conditions, including cardiovascular disease, diabetes, and mental health disorders [77,78]. Virtual care platforms may struggle to capture the full range of physical and cognitive assessments needed to effectively manage these conditions. In addition, older adults with HIV may face difficulties with digital literacy, which can hinder their ability to use virtual care tools effectively [79]. This is particularly true for older adults who may already experience cognitive decline or sensory impairments. The quality of the therapeutic relationship can also be affected in virtual settings as the absence of in-person interactions may reduce the ability of health care providers to observe subtle nonverbal cues and create a trusting environment [79]. Moreover, people living with HIV may have unique psychosocial challenges, such as stigma and social isolation, which can be exacerbated by virtual care [79]. Addressing these challenges requires the development of specialized, culturally competent virtual care models that incorporate both technological solutions and social support structures to ensure holistic, patient-centered care for older adults living with HIV [43].

Co-design is a valuable approach in developing and refining health interventions and services as it actively involves key players in the design process to ensure that the resulting solutions are highly relevant and effective for end users [80]. By engaging those who will ultimately benefit from these interventions, co-design fosters a deeper understanding of their needs, preferences, and sociocultural contexts, leading to more tailored and impactful outcomes [80,81]. By involving older adults in the design process, the resulting solutions will be more relevant and effective in addressing their unique health challenges and sociocultural contexts. *Co-design* promotes cultural sensitivity by incorporating the perspectives and experiences of diverse patient populations into the design and delivery of health services [82]. This is particularly important for older people living with HIV, who may face additional stigma and barriers underpinned by the intersection of sociocultural considerations and demographic factors such as age [83]. Addressing the gaps in the literature will ensure that the unique health challenges and sociocultural contexts of older people living with HIV are considered, leading to more effective and inclusive geriatric-HIV health care interventions [43].

Objectives

This paper reports on the co-design of a culturally appropriate virtual care model for older adults living with HIV using an experience-based co-design (EBCD) approach [84]. By involving older people living with HIV, this study sought to create a person-centered care model that enhances their health care experience and outcomes [85].

Methods

Study Design

In this community-based participatory research (CBPR), we used a qualitative EBCD methodology [84] that integrated participants as co-designers (details can be found in a published protocol [43]). CBPR is an approach that emphasizes collaboration between researchers and community members throughout the research process [86,87]. It aims to address community-identified needs and promote social change by integrating knowledge and action for social justice [86,88]. Co-design involves meaningful engagement of end users in the research design process, encompassing participation at every stage of the research and varying in intensity from passive involvement to highly active participation [84,89-91]. To achieve our research objective of co-designing a geriatric-HIV virtual care model, we adapted and used a variety of participatory EBCD methods, including interviews, focus groups, scenario design, and world cafés [84,92]. We briefly describe our methods in the following sections.

Ethical Considerations

Ethical considerations in CBPR include ensuring informed consent, where community members fully understand the research process and its implications [93]. This study was approved by the Sinai Health Research Ethics Board (approval 23-0106-E). All participants provided written informed consent to take part in the study, be audio recorded, and have anonymous quotations published. The participants were given the opportunity to have access to a translator and ask to clarify any concerns before signing the consent form. Participants in the focus groups and workshop were provided with refreshments and lunch and were reimbursed for their travel expenses. Participants in all phases also received an honorarium of $43 USD for their time, compensated at an hourly rate. Our study also emphasized community ownership of research findings, allowing the advisory committee to make decisions about how results are disseminated. All data shared with the advisory committee were anonymous.

Researcher Positionality

The research team (n=18) consisted of Canadian experts in co-design methodologies with diverse backgrounds and expertise in HIV research, equity-informed health care, health service research, and implementation science. This interdisciplinary team included PhD-trained researchers and clinicians specializing in geriatrics, family medicine, and infectious diseases, as well as occupational therapists, social workers,

health service administrators, and peer researchers with lived experience.

A community-based advisory team (n=10) was established before protocol development to ensure ongoing engagement throughout the project. This advisory team comprised knowledge users such as administrators from nonprofit organizations serving older persons living with HIV, including shelter staff, health care providers, charity organizations, and community centers. In addition, the advisory team included older individuals living with HIV and clinicians specializing in geriatric care. Regular individual meetings with principal investigators were conducted throughout the research process.

Recruitment and Participants

Our study sample was distinct from our advisory committee and included older individuals who self-identified as HIV positive; were aged ≥50 years; and resided in Ontario, Canada. We purposively sought diversity across various dimensions: (1) sex and gender, (2) age, (3) ethnicity and race, (4) socioeconomic status, (5) previous use of virtual geriatric care (ie, yes or no), (6) geographical location (rural vs urban), (7) time living with HIV, (8) non-English first language, and (9) level of educational attainment.

We used quota and purposive sampling to ensure a representative sample [94,95]. Our advisory team engaged in community outreach at relevant meetings and events to assist with recruitment, leveraging their networks and organizations. We also posted study flyers at community-based organizations, religious institutions, and culturally oriented events. Finally, our advisory team's websites and social media platforms were used to promote the study. All participants contacted the research coordinator to learn more about the study, and eligibility was confirmed based on the demographic characteristics to ensure alignment with our purposeful sampling technique.

Data Collection

Data collection occurred in 3 phases, which are outlined in this section and published in a previous protocol [43]. No deviations were made, aside from the fact that no focus groups occurred in phase 1. Phase 1 focused on understanding participants' needs and perspectives through semistructured interviews. The interviews were conducted by a trained research assistant with support from the primary author. The interviews focused on participants' experiences throughout their illness trajectory, current and anticipated needs for geriatric care, and priorities for virtual care. To ensure effective purposive quota sampling, demographic information was collected from participants at the time of recruitment, allowing us to determine whether they met the specific criteria for inclusion in the study. This demographic questionnaire captured diverse lived experiences and socioeconomic data relevant to this study's aims. Once participants were recruited, all interviews were audio recorded, professionally transcribed, and reviewed for accuracy by a research assistant. The narrative summaries of the interviews were created by the first author based on both the interviewer's notes and recollections and the transcripts of the interviews. These summaries were developed to provide a comprehensive overview of each interview, highlighting key themes, insights, and participant perspectives. Following their creation, the narrative summaries were shared with the research team and the advisory team via email and team meetings. The narrative summaries were included in the dataset for analysis and were considered part of the field and meeting notes, providing a synthesized account of key themes and insights from the interviews to support the interpretation and development of findings.

Before initiating the co-design process, we prioritized trust building to create a comfortable and inclusive environment for participants. To minimize power imbalances, we limited the number of researchers present, ensuring that the space felt less formal and more participant driven. In addition, we asked participants whom they wanted involved in the process, allowing them to shape the composition of the sessions. These efforts fostered a sense of ownership, encouraged open dialogue, and reinforced the collaborative nature of the co-design approach.

Phase 2 involved 2 focus groups and a workshop, all drawing on the same participants to apply an EBCD approach to develop a prototype for a culturally appropriate HIV virtual care model. The focus groups primarily aimed to gather in-depth qualitative data through structured discussions that allowed participants to share their experiences and insights about virtual care. In contrast, the workshop was more interactive and aimed at synthesizing the findings from the focus groups, where participants worked collaboratively to co-design and refine the elements of the HIV virtual care model. In preparation for phase 2, participants in the focus groups were provided with a detailed presentation summarizing the main needs identified in phase 1 before the workshop formally began. This presentation, based on comprehensive data collected during the initial phase, highlighted key insights into the unique challenges faced by older adults living with HIV. By sharing this information with the participants, we informed them about critical areas to address in the design process, ensuring that their feedback in the focus groups was grounded in the context of the identified needs. The focus groups and workshop were jointly led by a trained neutral facilitator and peer researcher who was trained through a structured program designed to build both facilitation and research skills. The training included an overview of the research objectives, ethical considerations, and techniques for creating a supportive and nonjudgmental environment for participants. The peer researcher was also trained in active listening, group dynamics, and how to manage sensitive topics, particularly those related to the experiences of older adults living with HIV. The workshop, guided by a modified 4D cycle of Appreciative Inquiry [96], aimed to identify strengths in existing care models and design principles for improved virtual care, accommodating both in-person and virtual attendance [97]. The modified 4D cycle of Appreciative Inquiry involves the phases of *Discovery*, *Dream*, *Design*, and *Destiny*, focusing on identifying strengths, envisioning an ideal future, cocreating actionable plans, and implementing sustained changes [96]. The *Definition* phase was also included for greater adaptability to specific contexts and goals. Discussions used visual tools such as concept mapping and word clouds [98] that were kept for analysis. All verbal focus group data were audio recorded and professionally

transcribed. Postworkshop reflections with the research and advisory teams guided phase 3.

In phase 3, there were 10 returning participants from phase 2 and 2 participants new to the study. The design principles for the culturally appropriate HIV virtual care model were developed using a world café format [99]. Small group discussions fostered dynamic exchanges, integrating feedback on proposed solutions and outlining recommendations for implementation and uptake. Facilitated by a peer researcher and supported by research team members, these sessions aimed to ensure participant engagement and generate robust data for ongoing project development [99]. Memos and reflexive notes were taken by the research team throughout the study. Transcripts and notes were anonymized by removing details that could lead to participant identification. No names of participants were used in the analysis or reporting of the results.

Data Analysis

The dataset comprised audio recordings, transcripts, field and meeting notes, reflexive notes, and physical artifacts designed by participants. The analysis team, consisting of research team members and select advisory committee members, conducted thematic content analysis [100,101]. The field and meeting notes were reviewed as part of the data analysis process, contributing to the identification of key themes. In phase 1, data analysis occurred concurrently with data collection to determine thematic saturation [102]. Thematic saturation was established when no new codes or concepts emerged from the data, indicating that additional interviews were unlikely to yield novel insights. Transcripts were coded using the NVivo software (version 14; QSR International) by the principal investigator and the research assistant. A codebook developed through inductive and deductive processes informed by the knowledge-to-action framework was used and refined iteratively through team meetings [103,104]. Physical artifacts created by participants during the workshop, including Post-it notes, drawings, and summary sheets, were also systematically analyzed and integrated into the overall data analysis process to enrich and deepen the understanding of participants' perspectives. Once scanned, the images were categorized based on themes and topics that emerged from the discussions. This allowed the research team to examine not only the content of the physical artifacts but also the ways in which participants visually expressed their ideas and emotions. The drawings and Post-it notes, for example, provided additional insights into participants' cognitive and emotional responses, offering a layer of qualitative data that complemented the verbal responses collected during the sessions. The images were analyzed using a thematic approach, identifying recurring patterns, symbols, and motifs that reflected key aspects of participants' experiences. Peer researchers were sent the preliminary findings 1 week ahead of the focus groups and workshop to enable thorough review and reflection before the subsequent phases. Themes were identified through comparative analysis and refined through team discussions, with a thematic map aiding in refining final themes to inform phase 2. Phase 2 and phase 3 continued with reflexive thematic content analysis to consolidate findings from the co-design workshop and refine themes for subsequent phases [101], including the world café discussions. The analysis refined the themes that emerged during these sessions and ensured that the feedback from the world café format was directly incorporated into the final design recommendations for the virtual care model. All researchers reviewed and endorsed the final version of the manuscript as coauthors to ensure the accuracy and completeness of the findings presented. A lay version of the manuscript was also sent to all members of the advisory team and participants of the study. Participants were invited to provide their feedback via email up to the point of manuscript submission.

Validation and Reliability Strategies

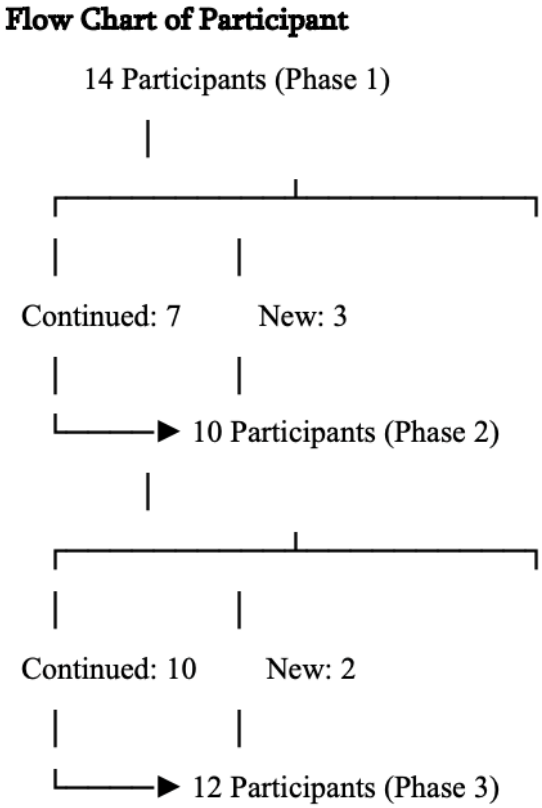
Rigor and trustworthiness in this study were ensured through several methodological strategies. First, we used a qualitative EBCD methodology, integrating participants as co-designers throughout the research process, which facilitated meaningful end-user engagement and ensured that participant perspectives were central to the study's design and outcomes. This approach was consistent with established principles of CBPR [86,105]. Second, we maintained methodological rigor by adhering strictly to our published protocol [43], which detailed the study procedures and ensured consistency in data collection and analysis methods. Our adherence to the protocol provided a structured framework for data collection and interpretation while remaining adaptable to participant and advisory committee input and emerging insights. This balance ensured that, while following a predefined approach, we also responded to and incorporated collaborators' feedback, thus enhancing the relevance and responsiveness of the methodology. Third, the study design included diverse methods that enhanced the richness and depth of the data collected and contributed to the triangulation of data sources [106,107]. The analysis team, consisting of research team members and selected advisory committee members, engaged in iterative coding and theme development supported by regular team meetings and peer review processes. This collaborative approach enhanced the reliability and credibility of the findings by ensuring that interpretations were grounded in the data and reflective of diverse perspectives [108]. Finally, transparency and reflexivity were maintained throughout the research process. That is, reflexive notes, memos, and detailed documentation of methodological decisions were kept, providing transparency in data interpretation and analysis. Regular interactions with the advisory team and participants, including sharing preliminary findings and seeking feedback, further strengthened the trustworthiness of this study by ensuring that the findings resonated with participant experiences and priorities [108].

Results

Overview

The research team conducted the co-design process between June 2023 and June 2024. In total, 19 unique participants took part in this study, with 14 (74%) participants in phase 1, a total of 10 (53%) participants in phase 2 (n=7, 70% from phase 1), and 12 (63%) participants in phase 3 (n=10, 83% from phases 1 and 2). Multimedia Appendix 1 outlines the participant characteristics, and Figure 1 shows a flowchart of the participants.

Figure 1. Flowchart of participants.



Participants co-designed a culturally appropriate and flexible HIV virtual care model that was rooted in empathy and addressed both the clinical, psychological, and social elements and the emotional aspects of aging with HIV. Phase 1 aimed to understand participants' needs regarding virtual care, revealing four key findings: (1) the desire for accessible and integrated health care; (2) the complications caused by sociocultural and demographic factors; (3) the benefits of peer support without replacing formal care; and (4) the potential of virtual care to address some barriers, although with caveats related to access and technology. Participants expressed a strong need for easily accessible, coordinated care, particularly for older adults living with HIV who face complex, overlapping health and social challenges. They highlighted the difficulties of navigating the health care system, noting that virtual care could reduce barriers but only if it accounts for digital literacy and technology access. In addition, while peer support was valued, participants cautioned against overreliance on it, emphasizing that virtual care should supplement, not replace, formal services. They also pointed out that virtual care could improve access, especially for individuals facing logistical barriers, but it must be designed with flexibility and continuity in mind. In phase 2, the focus shifted to ideating solutions for the virtual geriatric-HIV care model. Participants suggested a multimodal care approach encompassing various forms of care delivery (eg, in-person and virtual visits and digital health records). They also stressed the importance of education and clear communication regarding technology use, ensuring that older adults are informed and confident in using virtual platforms. Privacy and data security were significant concerns, with participants requesting transparency about who accesses their information and how it is used. In addition, participants emphasized the need for reminder systems to help maintain continuity of care and reduce missed appointments. These solutions were aimed at improving both the practicality and trustworthiness of virtual care for older adults living with HIV. Phase 3 focused on co-designing actionable plans, with participants emphasizing the importance of linguistic and cultural equity in virtual care. They stressed that care models should be inclusive of diverse populations, particularly those who face language barriers. Participants highlighted the need for culturally appropriate services to ensure equitable care, pointing out that older adults from different cultural backgrounds often experience compounded difficulties when accessing health care services. The plans developed during this phase aimed to refine and prioritize these solutions to create a more inclusive and effective virtual care model for older adults living with HIV. Throughout the following sections, we illustrate the findings using quotes labeled with participant ID, race or ethnicity, gender, and age.

Phase 1: Understanding Participants' Needs Regarding Virtual Care

We identified the following key and interrelated insights: (1) older adults living with HIV want accessible and integrated care, (2) intersecting sociocultural and demographic factors complicate receipt of care for older adults living with HIV, (3) peer support is beneficial but should not replace formal health care services and supports, and (4) virtual care can address barriers to care with access caveats.

Older Adults Living With HIV Want Accessible and Integrated Care

Participants emphasized the need for health care systems to prioritize services that are easily accessible and tailored to the unique health challenges faced by older individuals living with HIV. Participants noted that they often had complex, overlapping medical and social needs that required coordinated care across multiple disciplines. While most participants described positive experiences with previous care, they also discussed challenges obtaining information and reaching the appropriate care provider. Several participants specifically recounted situations in which accessing services and support felt very difficult. One participant shared the following:

> *You can go to a family doc[tor] if you have one and that is so great—again, if you have one. But the issue becomes when the doc refers you to one specialist who refers you to another then back to your family doc and you repeat because no one knows how to care for you.* [Interview participant 11; man; White individual; aged 52 years]

Therefore, participants described desiring a health care system with the ability to easily access services from care providers.

Intersecting Sociocultural and Demographic Factors Complicate Receipt of Care for Older Adults Living With HIV

Participants highlighted that they often faced systemic barriers that hindered their ability to access health care services. These challenges were exacerbated by age-related biases within the system, which was not designed to fully address or accommodate their unique needs. Participants highlighted that those with lower incomes may experience additional obstacles, including lower access to resources; heightened stigma; and greater socioeconomic disadvantages that impact their health, such as the inability to purchase both food and medication. Participants agreed that virtual care has the potential to alleviate some of these systemic barriers by providing more accessible and flexible health care options. However, for virtual care to be an effective solution, it must address the challenges associated with digital literacy and technology access. One participant pointed out the following:

> *The level that we're to access the services at, is a bit crazy and some might have challenges. Like this morning, I tried to go on the touch screen, and I can't get it on to use my phone. And then I'm seeing that there is no Internet service.* [Interview participant 5; man; Black individual; aged 56 years]

Moreover, to ensure that virtual care is equitable, it is crucial that health care providers are aware of and address the specific needs and preferences of older adults. For instance, participants noted that they preferred a larger screen for virtual consultations, such as a tablet, over smaller devices such as phones.

Peer Support Is Beneficial but Should Not Replace Formal Health Care Services and Supports

Support systems among peers were noted as critical in the management of chronic conditions among older people living with HIV. Peer support groups offered participants emotional and practical guidance from individuals who shared similar experiences, fostering a sense of community and reducing feelings of isolation. Peer interactions were also thought to provide valuable insights into managing HIV and navigating the health care system as an older person as they offered lived experiences and practical advice that complemented professional care. One participant noted the positive impact of peer support, stating the following:

> *I found that connecting with others who are dealing with the same condition really helped me cope better and stay informed about my health.* [Interview participant 3; man; South Asian individual; aged 82 years]

In the context of peer support, health management and the pursuit of independence emerged as critical themes in the experiences of older adults living with HIV, particularly within a virtual care framework. Participants relied heavily on peer support to navigate the health care system, highlighting the need for virtual care models that offer tailored resources and programs to address their unique needs. One participant emphasized the value of community-based programs, noting the following:

> *First of all, it's a safe place for LGBTQ members. And it's a positive place for me because it's 50 plus. Basically, it focuses on age and being gay and what more can we want. Canada is great—it won't discriminate by colour in medical settings, but sometimes we can't find what we need.* [Interview participant 11; man; White individual; aged 52 years]

Despite the availability of such programs, challenges remained, particularly in accessing mental health services. The participants wished for more targeted online resources for aging, stress, or depression beyond medication. Furthermore, virtual care could enhance independence in health management by facilitating regular interactions with health care providers through digital platforms. Participants expressed a need for frequent virtual check-ins, such as weekly phone calls, to better monitor and manage their health conditions and maintain a sense of connection and support.

Participants noted that, while peer support was invaluable, they were cautious that their reliance on individuals with lived experience was due to systemic gaps in health care provision and the lack of formal support structures for addressing the complex interplay of health and social circumstances affecting older adults living with HIV. One participant remarked the following:

> *If I get sick, then I lose my job, and then the chance of losing my housing rise exponentially and so basically I have to have friends who have a house or know someone who is hiring to pay for medicine.* [Interview participant 1; woman; White individual; aged 57 years]

The same participant shared the following:

> *If you have money, you'll get the healthcare in the time that you need it. And if you don't, well, then good luck.*

Another participant shared the following:

> *When you're first diagnosed, all you're focused on is survival. You aren't trying to find services; you meet people who tell you what services exist. But the healthcare system, the focus isn't on us, who've been around for a long time, and now we need different supports in place than we did 20 years ago. So, we're stuck with the same services.* [Interview participant 2; woman; First Nations individual; aged 57 years]

Virtual Care Can Address Barriers to Care With Access Caveats

Participants noted that the integration of virtual care has the potential to alleviate some of these challenges by offering more flexible and accessible health care options, thereby significantly reducing the logistical challenges associated with in-person visits, as noted by a participant who appreciated the time savings:

> *I had a bone density scan...when I had it, he wanted to make an appointment, and I had the choice between virtual or going down to the office. I saved like 4 hours.* [Interview participant 12; man; First Nations individual; aged 51 years]

Another participant shared the following:

> *When the place is very cold and you got to catch two bus, three busses to meet your doctor, and pay for it each time, we can do virtual. Then when it's nice outside whether we don't mind going and spending the time and go out and see.* [Interview participant 10; woman; mixed heritage; aged 54 years]

However, one participant wisely pointed out the following:

> *But it has to be only when minor consultation is required...it should [not] replace the face-to-face consultation.* [Interview participant 9; man; White individual; aged 59 years]

Participants emphasized the importance of effective communication and continuity of care within the virtual geriatric-HIV care model. They expressed a strong desire for seamless interactions among care providers to prevent disruptions in their treatment and communication. This continuity was described as crucial for maintaining the stability of their HIV management, including regular blood work and medication adjustments. Although participants generally reported positive interactions with health care providers, challenges arose in distinguishing between symptoms related to HIV and those associated with normal aging. One participant remarked the following:

> *Sometimes it's hard to tell whether a symptom is due to the HIV or just getting older. Clear communication and understanding from my healthcare provider are essential to figuring out what's really going on.* [Interview participant 9; man; White individual; aged 59 years]

To address these complexities, participants recommended that novel care models integrate mechanisms for clear, ongoing communication and coordination among health care providers, ensuring that specialized care is tailored to meet the intricate needs of older adults living with HIV.

Phase 2: Ideating Solutions

Participants suggested improvements to current care models for older adults living with HIV and proposed their wishes regarding virtual care through a series of activities.

Importance of Flexible Care Delivery

Participants highlighted the importance of a multimodal care approach to enhance coordinated care and access to health information for older adults living with HIV. One participant explained the following:

> *...we need it [virtual care] to be offered in a lot of ways, so not just phones because not everyone has one, not everyone has a cell phone, maybe computer, so offer a lot of ways like a phone, cellphone, tablet, computer.* [Phase-2 participant 1; man; White individual; aged 56 years]

Participants also noted that the approach should integrate various modes of care delivery, such as in-person consultations, virtual visits, and digital health records, to create a more seamless, continuous, and comprehensive health care experience. One participant shared the following:

> *...you can't just have it be one approach, sometimes you need to come in, sometimes I might need a virtual visit, sometimes I might want to send information, we need options.* [Phase-2 participant 4; woman; Black individual; aged 71 years]

Education and Communication Regarding Technology Use and Privacy

The need for clear communication regarding new technologies was a recurring theme among participants. They emphasized that educational demonstrations are essential in helping patients understand and effectively use new technological tools. One participant said the following:

> *...the worst thing someone can do is if you assume we know what we are doing with technology, so it doesn't have to be a doctor but some instructions. I sometimes use Zoom but then they say go online and upload, I have no idea how to do that.* [Phase-2 participant 6; man; White individual; aged 62 years]

Transparency regarding technology use in virtual care emerged as a crucial concern for participants. Participants emphasized the need for clear communication about who accesses patient information, how it is used, and the procedures for withdrawing consent. In the context of virtual care, the digital environment can heighten concerns about privacy and confidentiality, making it imperative to establish transparent practices in data privacy and security. Participants indicated that, in virtual settings, where information is often shared and stored electronically, patients may feel more vulnerable and uncertain about their data. Therefore, ensuring that patients are well informed about the handling of their health information is essential for building trust and fostering a sense of comfort with using virtual care services. One participant explained the following:

> *...sometimes they think we are old, so we don't understand, but I know about technology and so I want to know who is going to look at my stuff in order to trust you. If I don't know, I won't trust this [virtual care].* [Phase-2 participant 8; woman; Indian-Caribbean individual; aged 54 years]

Participants also underscored the need for transparency regarding the environment in which virtual care appointments are conducted. They expressed concerns about ensuring that these appointments occur in private and secure settings to protect patient confidentiality. Specifically, participants highlighted the importance of knowing whether providers are conducting virtual appointments from locations that guarantee privacy and maintain the integrity of sensitive health information:

> *I don't want the doctor to be like talking about HIV with their friends in the background or at Starbucks.* [Phase-2 participant 8; woman; Southeast Asian individual; aged 70 years]

Reminder Systems

Participants suggested the implementation of reminder systems for appointments and follow-ups to ensure continuity of care and reduce missed appointments. Such systems can play a crucial role in maintaining regular contact between patients and health care providers, thereby improving adherence to treatment plans and enhancing health outcomes. By leveraging technology to provide timely reminders, health care systems can support patients in managing their care more effectively and reduce the risk of lapses in treatment. One participant shared the following:

> *...maybe it's age, maybe it's life, but anything to make it easier is good. Remind me when to log in, when to book an appointment, what to do, who to see, even the name, just anything to make life easier when it comes to doctors because sometimes, I see so many and now maybe a geriatrician.* [Phase-2 participant 4; woman; Black individual; aged 71 years]

Phase 3: Designing the Action Plan

Overview

Phase 3 entailed participants co-designing and planning the necessary actions and initiatives to achieve the desired future for virtual geriatric-HIV care as outlined in phase 2. Participants concluded phase 3 by refining the ideas and virtual geriatric care models co-designed in phase 2 and phase 3. This involved prioritizing linguistic and cultural equity.

Linguistic and Cultural Equity

Participants highlighted the need for the virtual geriatric care model to incorporate linguistic and cultural considerations to ensure equitable care for the population of older adults living with HIV. Participants shared stories exemplifying the diversity of the geriatric-HIV community (inclusive of several underserved populations) and emphasized the importance of increasing services to these populations. One participant stated the following:

> *...not everyone speaks English, especially when they come here and so we need to give back and help them. It's like if you dropped someone in their countries, it's like worse than a maze.* [Phase-3 participant 11; man; White individual; aged 69 years]

Participants noted that there needs to be health- and technology-related information in multiple languages and that service providers should make translation services known to their patients. Participants also suggested that health care providers offer education that addresses many of the culturally related myths and concerns of individuals about aging with HIV. One participant said the following:

> *...doctors and nurses need to look at who they are serving that is different than them. Not everyone understands HIV is not a life sentence so tell certain people that. Or maybe someone eats food that is different, not everyone understands quinoa. Maybe recommend cultural food.* [Phase-3 participant 1; woman; Black individual; aged 51 years]

Consider Affordability

Participants were broadly aware of the fact that the cost of technology could be a barrier to many older adults living with HIV. Participants highly recommended that geriatric-HIV virtual care models be made more accessible by offering subsidized internet access or technological devices. Participants also shared that some patients may have never used technology before virtual care due to high costs and, therefore, may need training on how to best use the technology for care. Participants recognized that health care providers may not have the capacity to teach their clients, and as such, someone else in the health care organization should. For instance, participants suggested that perhaps calling a patient on a landline or inviting the patient to in-person training could help support their use of virtual modalities for health care in the future. One participant described the following:

> *...maybe invite someone to a place in person just even once or call them and walk them through something like how to turn their video on or how to send a text to a doctor.* [Phase-3 participant 17; woman; Black individual; aged 54 years]

Participants emphasized the importance of early education on technology, as previously underscored in the other phases of the study.

Participants expressed significant concern about the high cost of HIV-related medications, particularly for those aged <65 years, who are ineligible for government-funded drug coverage in Ontario. Some participants felt that health care providers were not advocating enough for the use of lower-cost alternatives such as generic medications, which could be just as effective as brand-name drugs. For instance, one patient mentioned the following:

> *It [medication] can be the choice between food, rent or medications at times, so I want someone who maybe says a cheaper supplement or maybe what medication I can maybe take a different type of. Anything to keep the cost low.* [Phase-3 participant 9; woman; Asian individual; aged 70 years]

Participants indicated that a more proactive stance from providers in promoting cost-effective treatment options, both for HIV and age-related conditions, would improve their overall care experience.

Person-Centered Care for Age-Related Health Issues

Participants highlighted the importance of having health care providers who understood the dual challenges of aging and living with HIV. They emphasized that a key feature of an effective geriatric virtual care program would be the integration of expertise in both areas. This holistic approach was seen as essential for managing the complexities of age-related conditions alongside HIV, with many noting that current segmented specialist care often leads to fragmented treatment, polypharmacy, and inconsistent care coordination. An integrated model was viewed as a potential solution to improve care continuity and reduce these challenges.

Participants emphasized the importance of incorporating cognitive health screening and related support services into the virtual care model. This integration would address current gaps in HIV care by ensuring that the program not only focuses on medical treatment but also considers the emotional aspects of aging with HIV. Current gaps within HIV care were re-emphasized, and participants noted that, for uptake of the geriatric-HIV virtual care model to occur, participants needed to "feel as if the doctor or care professionals care of us as a person, not just a billable code or something, but to really understand how much [older adults living with HIV] struggle" (phase-3 participant 8; man; White individual; aged 67 years).

Patient Resilience and Support

Participants emphasized a virtual geriatric-HIV model of care that incorporates support networks to enhance patient resilience to living with HIV in older age. The involvement of trusted individuals such as existing care providers, friends, or service agencies was deemed essential to care by participants. This holistic support system helps patients navigate the challenges of their health journey, promotes a sense of connection and belonging, and strengthens their overall well-being. By embedding these peer support networks within the virtual care model, the program can provide a more robust and empathetic approach to patient care. One participant described the following:

> *I'd trust anything that [Community Agency] recommended, so if you bring your services to them so that they encourage them, I'll trust that it would help me.* [Phase-3 participant 4; woman; White individual; aged 58 years]

Moreover, participants emphasized that a virtual geriatric-HIV model of care should include features such as virtual support groups, peer mentoring programs, and online forums where individuals with similar experiences can connect and share insights. In addition, incorporating the peer mentorship system could allow patients to receive guidance and encouragement from those who have successfully managed similar health conditions.

Discussion

Principal Findings

This study was conducted to co-design a geriatric-HIV virtual care model for older adults living with HIV. The lack of tailored interventions that consider this population's specific needs and preferences necessitated a co-design approach. Using EBCD methodology [84,85,92], we engaged older adults living with HIV to collaboratively codevelop a virtual care model. Through interviews, focus groups, and a co-design workshop, participants identified key needs such as robust support systems, health management resources, and effective communication strategies. They cocreated and refined solutions including multimodal care options, technology transparency recommendations, and reminder systems.

Using co-design methodologies ensured that the voices and experiences of older adults living with HIV were central to the development and evaluation of the virtual care models. This inclusive, user-centered approach fosters solutions that are more likely to be accepted and effective as they are directly informed by the end users' needs and preferences. Engaging participants in the research process empowers them and enhances the relevance and applicability of the findings. A strength of this study was the inclusion of many English-as-a-second-language individuals and immigrants to Canada, who are typically underrepresented in research [109]. This diverse participant pool enhanced the transferability of the findings, ensuring that the developed virtual care models were relevant and accessible to a broader range of older adults living with HIV. By actively including these often-excluded groups, this study addresses potential disparities in health care access and provides insights into the unique needs and challenges faced by these populations. Thus, this research provides practical insights into the real-world application and potential barriers to adoption of virtual care models. This dual focus ensures that this study not only explores the theoretical potential of the interventions but also addresses the pragmatic challenges that might arise during implementation. As a result, the findings are highly relevant for informing policy and practice, ultimately contributing to the advancement of care for older adults living with HIV.

The findings presented in this paper have been echoed in various other studies with other populations. Research among rural and urban adults has shown high satisfaction with virtual care [110]. In addition, studies on virtual care for substance use indicate its ability to bridge service gaps, particularly in rural areas, reinforcing the importance of multimodal care options and effective communication [111]. Canadian case studies further demonstrate the necessity of involving users in the development of virtual care tools to ensure their effectiveness and transparency [112].

Our findings underscore the importance of integrating cultural and linguistic equity into virtual care models, which has substantial implications for enhancing health care accessibility and outcomes for older adults living with HIV. Research has shown that sociocultural stigma can severely impact health care engagement and access [36,113], highlighting the importance of designing solutions that are sensitive to these issues.

Involving older adults living with HIV in co-designing care solutions is essential to ensure that their needs are met [90]. Engaging older adults living with HIV in the development process ensures that the solutions are tailored to their specific needs and the challenges they face, thus enhancing their effectiveness and acceptance [90]. This study's findings have important implications for health policy, particularly regarding the integration of virtual care models for older persons living with HIV. Existing reviews have found that virtual geriatric-HIV programs are limited [17] despite their proposed benefits. Policy makers should consider the advantages of virtual care in addressing the unique needs of this population, especially in contexts in which traditional health care access is limited. Our study highlights that participants perceive virtual care to be feasible and acceptable for meeting their geriatric-HIV care needs, suggesting that policies should support the expansion of telehealth services and funding for technological infrastructure. In addition, policies should focus on inclusivity by ensuring that virtual care models are accessible to non-English speakers and diverse cultural groups who may have differing views on HIV care [50,114]. This could involve implementing language support services and cultural competency training for health care providers to enhance the effectiveness of virtual interventions. This could also involve funding geriatric health care settings such that they can hire multilingual staff, partner with translation services, and provide localized materials.

In practice, this study underscores the importance of adopting virtual care models as a viable alternative or complement to traditional in-person geriatric consultations for older persons living with HIV [67]. Health care providers should be encouraged to incorporate these models into their service offerings, with attention to user feedback and the co-design process highlighted in this study. Research emphasizes the need for health care professionals to collaborate with patients in co-design methodologies [115] to enhance patient participation in health care services. Projects involving creative workshops have shown that health care professionals can gain valuable insights and training through active participation in co-design processes [116]. In addition, discussions on EBCD underscore the necessity of equipping health care professionals with the tools and knowledge to effectively engage in patient-centered design and improvement initiatives [117]. Overall, training health care providers in these collaborative approaches is essential for fostering meaningful partnerships with patients in health care settings. For example, health care settings could employ staff trained in the use of virtual platforms and equipped with resources to support patients in navigating these technologies. Optimal methods for training older adults to use technology involve a combination of hands-on practice, clear instructions, and personalized support tailored to individual needs and cognitive abilities [118]. Effective approaches often include guided action and attention training, addressing fears, building confidence, and providing ongoing support, with a focus on practical, relevant applications that motivate learners [118,119]. Moreover, the integration of language support and culturally sensitive practices is crucial to ensure that virtual care is accessible to diverse populations. Other studies of virtual care have found that features such as multilingual support, screen readers, and voice commands can make virtual care more accessible to diverse patient populations [120]. Moreover, integrating peers into virtual care models can be achieved through comprehensive support, training, and collaboration across various providers [121].

To test the virtual geriatric care model, future research should use randomized controlled trials comparing virtual geriatric-HIV care with traditional care in terms of patient satisfaction, health outcomes, and cost-effectiveness. Specific technological platforms and features should also be evaluated to determine their effectiveness in addressing the unique needs of culturally diverse older adults living with HIV. In addition, implementation studies should examine barriers and facilitators in various care settings to optimize and scale the model effectively. Future research should build on the findings of this study by exploring the long-term impacts of virtual care models on health outcomes and quality of life for older persons living with HIV. Longitudinal studies could provide insights into the sustainability and efficacy of these interventions over time. In addition, research should investigate the effectiveness of different technological platforms and features in meeting the needs of culturally diverse older persons living with HIV. Existing comparative studies have not assessed virtual care against traditional models in the context of HIV. In addition, issues of equity and access must be addressed to avoid further disadvantaging vulnerable individuals who may lack access to necessary technology or resources. Furthermore, exploring the barriers to and facilitators of implementation in various settings, such as acute and community care, can offer valuable information for refining and scaling virtual geriatric-HIV care models. Finally, future research should continue to address the needs of non-English speakers and other underserved groups, ensuring that virtual care remains inclusive and equitable.

Limitations

This study, while innovative in its co-design approach to developing a virtual care model for older adults living with HIV, has noted limitations. First, the sample size was relatively small and geographically limited to a dense urban center, which may impact the transferability of the findings. Most participants were already engaged with technology, which may not accurately represent the broader population of older adults living with HIV who have varying levels of technological proficiency. This existing familiarity could have positively biased the feedback on virtual care models as these participants, who were actively seeking health care services, might be more comfortable and adept with digital tools. This may not be representative of others who avoid health care and may struggle more with virtual care platforms. Consequently, the findings may overestimate the acceptability of such interventions for the entire demographic, potentially overlooking significant barriers experienced by those less familiar or comfortable with technology. Future research should aim to include a more diverse sample in terms of technological experience to ensure that the developed care models are inclusive and accessible to all older adults living with HIV. Finally, geriatric-HIV care providers and community agencies were not engaged as formal participants in the co-design phases. Their exclusion may have limited insights into the clinical and service delivery perspectives necessary for developing a more holistic and integrated care model. Engaging

these stakeholders could have addressed gaps related to continuity of care and support services. Future iterations of this model would benefit from the involvement of health care providers and community organizations to ensure that the virtual care system is aligned with existing services and adequately meets the needs of older adults living with HIV.

Conclusions

This study highlights the crucial need for socioculturally appropriate virtual care models for older persons living with HIV. Using an EBCD approach, we engaged older persons living with HIV in co-designing a virtual care model that addresses their unique needs and challenges. The findings reveal several key areas for improvement in virtual geriatric-HIV care, including the necessity for robust support systems, effective communication strategies, and tailored health management resources. Participants identified critical needs such as the integration of multimodal care options, transparency in technology use, and the implementation of reminder systems to enhance patient engagement and continuity of care. They also emphasized the importance of linguistic and cultural equity, recommending that virtual care models incorporate diverse language options and cultural considerations to ensure accessibility and relevance. In addition, affordability was highlighted as a significant barrier, with participants advocating for subsidized technology and training to overcome economic challenges. Policy makers should consider the advantages of virtual care for older persons living with HIV and support the expansion of telehealth services and technological infrastructure. Future research should explore the long-term impacts of virtual care on health outcomes and quality of life on older persons living with HIV as well as evaluating different technological platforms and features. By embracing the diverse voices of older persons living with HIV, this study paves the way for a more inclusive, innovative, and equitable virtual future of care.

Conflicts of Interest

KMK is an Associate Editor for JMIR Rehabilitation and Assistive Technologies.

Multimedia Appendix 1
Demographic information for participants in phases 1, 2, and 3 (N=19).
[DOCX File , 41 KB - aging_v8i1e67122_app1.docx]

References

1. He W, Kinsella K. Global aging in the new millennium. In: Sokolovsky J, editor. The Cultural Context of Aging: Worldwide Perspectives. Westport, CT: Praeger/ABC-CLIO; 2020:27-46.
2. Autenrieth CS, Beck EJ, Stelzle D, Mallouris C, Mahy M, Ghys P. Global and regional trends of people living with HIV aged 50 and over: estimates and projections for 2000-2020. PLoS One 2018 Nov 29;13(11):e0207005 [FREE Full text] [doi: 10.1371/journal.pone.0207005] [Medline: 30496302]
3. Sapuła M, Suchacz M, Załęski A, Wiercińska-Drapało A. Impact of combined antiretroviral therapy on metabolic syndrome components in adult people living with HIV: a literature review. Viruses 2022 Jan 11;14(1):122 [FREE Full text] [doi: 10.3390/v14010122] [Medline: 35062326]
4. Grosso TM, Hernández-Sánchez D, Dragovic G, Vasylyev M, Saumoy M, Blanco JR, et al. Identifying the needs of older people living with HIV (≥ 50 years old) from multiple centres over the world: a descriptive analysis. AIDS Res Ther 2023 Feb 12;20(1):10 [FREE Full text] [doi: 10.1186/s12981-022-00488-7] [Medline: 36782210]
5. Guaraldi G, Milic J, Mussini C. Aging with HIV. Curr HIV/AIDS Rep 2019 Dec 10;16(6):475-481. [doi: 10.1007/s11904-019-00464-3] [Medline: 31823125]
6. Wing EJ. HIV and aging. Int J Infect Dis 2016 Dec;53:61-68 [FREE Full text] [doi: 10.1016/j.ijid.2016.10.004] [Medline: 27756678]
7. Onen NF, Overton ET, Seyfried W, Stumm ER, Snell M, Mondy K, et al. Aging and HIV infection: a comparison between older HIV-infected persons and the general population. HIV Clin Trials 2010 Jan 06;11(2):100-109. [doi: 10.1310/hct1102-100] [Medline: 20542846]
8. Aberg JA. Aging and HIV infection: focus on cardiovascular disease risk. Top Antivir Med 2020 Jan;27(4):102-105 [FREE Full text] [Medline: 32224501]
9. Martin-Iguacel R, Llibre JM, Friis-Moller N. Risk of cardiovascular disease in an aging HIV population: where are we now? Curr HIV/AIDS Rep 2015 Dec 30;12(4):375-387. [doi: 10.1007/s11904-015-0284-6] [Medline: 26423407]
10. Medapalli RK, Parikh CR, Gordon K, Brown ST, Butt AA, Gibert CL, et al. Comorbid diabetes and the risk of progressive chronic kidney disease in HIV-infected adults: data from the Veterans Aging Cohort Study. J Acquir Immune Defic Syndr 2012 Aug 01;60(4):393-399 [FREE Full text] [doi: 10.1097/QAI.0b013e31825b70d9] [Medline: 22592587]
11. Biver E. Osteoporosis and HIV Infection. Calcif Tissue Int 2022 May 30;110(5):624-640 [FREE Full text] [doi: 10.1007/s00223-022-00946-4] [Medline: 35098324]
12. Brown TT. HIV: an underrecognized secondary cause of osteoporosis? J Bone Miner Res 2013 Jun;28(6):1256-1258 [FREE Full text] [doi: 10.1002/jbmr.1967] [Medline: 23630095]

13. Cotter AG, Mallon PW. HIV infection and bone disease: implications for an aging population. Sex Health 2011;8(4):493. [doi: 10.1071/sh11014]
14. Frain J. Considering the impact of aging when caring for and treating adults with HIV. J Gerontol Nurs 2020 Apr 01;46(4):31-40. [doi: 10.3928/00989134-20200304-02] [Medline: 32219455]
15. Cherry ST, Tyagi K, Bailey J, Reagan B, Jacuinde V, Soto J, et al. More to come: perspectives and lived experiences of adults aging with HIV. J Gay Lesbian Soc Serv 2021 Jul 01;34(2):177-206. [doi: 10.1080/10538720.2021.1946669]
16. Cox LE, Brennan-Ing M. Medical, social and supportive services for older adults with HIV. Interdiscip Top Gerontol Geriatr 2017;42:204-221. [doi: 10.1159/000448563] [Medline: 27875834]
17. Kokorelias KM, Grosse A, Zhabokritsky A, Sirisegaram L. Understanding geriatric models of care for older adults living with HIV: a scoping review and qualitative analysis. BMC Geriatr 2023 Jul 08;23(1):417 [FREE Full text] [doi: 10.1186/s12877-023-04114-7] [Medline: 37422631]
18. Kokorelias KM, Grosse A, Zhabokritsky A, Walmsley SL, Sirisegaram L. Exploring implementation considerations for geriatric-HIV clinics: a secondary analysis from a scoping review on HIV models of geriatric care. HIV Med 2024 Jan 04;25(1):38-59. [doi: 10.1111/hiv.13549] [Medline: 37794607]
19. Brothers TD, Rockwood K. Frailty: a new vulnerability indicator in people aging with HIV. Eur Geriatr Med 2019 Apr 10;10(2):219-226. [doi: 10.1007/s41999-018-0143-2] [Medline: 34652747]
20. Kalichman SC, Heckman T, Kochman A, Sikkema K, Bergholte J. Depression and thoughts of suicide among middle-aged and older persons living with HIV-AIDS. Psychiatr Serv 2000 Jul;51(7):903-907. [doi: 10.1176/appi.ps.51.7.903] [Medline: 10875956]
21. Halkitis PN, Krause KD, Vieira DL. Mental health, psychosocial challenges and resilience in older adults living with HIV. Interdiscip Top Gerontol Geriatr 2017;42:187-203. [doi: 10.1159/000448564] [Medline: 27875833]
22. Mavandadi S, Zanjani F, Ten Have TR, Oslin DW. Psychological well-being among individuals aging with HIV: the value of social relationships. J Acquir Immune Defic Syndr 2009 May 01;51(1):91-98 [FREE Full text] [doi: 10.1097/QAI.0b013e318199069b] [Medline: 19282781]
23. Brandt CP, Sheppard DP, Zvolensky MJ, Morgan EE, Atkinson JH, Woods SP. Does age influence the frequency of anxiety symptoms and disorders in HIV disease? J HIV/AIDS Soc Serv 2016 Jun 28;15(4):380-403. [doi: 10.1080/15381501.2016.1189865]
24. Basu S, Chwastiak LA, Bruce RD. Clinical management of depression and anxiety in HIV-infected adults. AIDS 2005 Dec 02;19(18):2057-2067. [doi: 10.1097/01.aids.0000182518.84407.32] [Medline: 16284454]
25. Deng L, Zhang X, Gao Y, Turner D, Qian F, Lu H, et al. Association of HIV infection and cognitive impairment in older adults: a meta-analysis. Ageing Res Rev 2021 Jul;68:101310 [FREE Full text] [doi: 10.1016/j.arr.2021.101310] [Medline: 33640473]
26. Alford K, Vera JH. Cognitive impairment in people living with HIV in the ART era: a review. Br Med Bull 2018 Sep 01;127(1):55-68. [doi: 10.1093/bmb/ldy019] [Medline: 29868901]
27. Hosaka KR, Greene M, Premeaux TA, Javandel S, Allen IE, Ndhlovu LC, et al. Geriatric syndromes in older adults living with HIV and cognitive impairment. J Am Geriatr Soc 2019 Sep 26;67(9):1913-1916 [FREE Full text] [doi: 10.1111/jgs.16034] [Medline: 31241764]
28. Grov C, Golub SA, Parsons JT, Brennan M, Karpiak SE. Loneliness and HIV-related stigma explain depression among older HIV-positive adults. AIDS Care 2010 May 16;22(5):630-639. [doi: 10.1080/09540120903280901] [Medline: 20401765]
29. Brown MJ, Adeagbo O. HIV and aging: double stigma. Curr Epidemiol Rep 2021 Mar 12;8(2):72-78 [FREE Full text] [doi: 10.1007/s40471-021-00265-6] [Medline: 33728256]
30. Brennan-Ing M. Diversity, stigma, and social integration among older adults with HIV. Eur Geriatr Med 2019 Apr 07;10(2):239-246. [doi: 10.1007/s41999-018-0142-3] [Medline: 34652745]
31. DeMarco RF, Brennan-Ing M, Sprague C, Brown SM. Ageism, aging and HIV: community responses to prevention, treatment, care and support. Interdiscip Top Gerontol Geriatr 2017;42:234-239. [doi: 10.1159/000448567] [Medline: 27875836]
32. Emlet CA. "You're awfully old to have this disease": experiences of stigma and ageism in adults 50 years and older living with HIV/AIDS. Gerontologist 2006 Dec;46(6):781-790. [doi: 10.1093/geront/46.6.781] [Medline: 17169933]
33. Wallach I, Brotman S. Gaps in health and social care services to older adults living with HIV: a qualitative study on the perspectives of older adults and service providers. Nurs Care Open Access J 2019 Jan 21;6(1). [doi: 10.15406/ncoaj.2019.06.00179]
34. O'Neill G, Barry PP. Training physicians in geriatric care: responding to critical care. Public Policy Aging Rep 2003;13(2):17-21 [FREE Full text] [doi: 10.1093/ppar/13.2.17]
35. Harris J. Geriatric trends facing nursing with the growing aging. Crit Care Nurs Clin North Am 2019 Jun;31(2):211-224. [doi: 10.1016/j.cnc.2019.02.007] [Medline: 31047094]
36. Airhihenbuwa CO, Ford CL, Iwelunmor JI. Why culture matters in health interventions: lessons from HIV/AIDS stigma and NCDs. Health Educ Behav 2014 Feb 17;41(1):78-84. [doi: 10.1177/1090198113487199] [Medline: 23685666]

37. Arnold EA, Rebchook GM, Kegeles SM. 'Triply cursed': racism, homophobia and HIV-related stigma are barriers to regular HIV testing, treatment adherence and disclosure among young Black gay men. Cult Health Sex 2014 Jun 02;16(6):710-722 [FREE Full text] [doi: 10.1080/13691058.2014.905706] [Medline: 24784224]
38. Faller J, Marcon SS. Praticas socioculturais e de cuidado a saude de idosos em diferentes etnias. Esc Anna Nery 2013 Sep;17(3):512-519. [doi: 10.1590/S1414-81452013000300015]
39. Weßel M. Feminist approach to geriatric care: comprehensive geriatric assessment, diversity and intersectionality. Med Health Care Philos 2022 Mar 16;25(1):87-97 [FREE Full text] [doi: 10.1007/s11019-021-10052-1] [Medline: 34529218]
40. Holden K, McGregor B, Thandi P, Fresh E, Sheats K, Belton A, et al. Toward culturally centered integrative care for addressing mental health disparities among ethnic minorities. Psychol Serv 2014 Nov;11(4):357-368 [FREE Full text] [doi: 10.1037/a0038122] [Medline: 25383991]
41. Vu GT, Tran BX, Hoang CL, Hall BJ, Phan HT, Ha GH, et al. Global research on quality of life of patients with HIV/AIDS: is it socio-culturally addressed? (GAPRESEARCH). Int J Environ Res Public Health 2020 Mar 23;17(6):2127 [FREE Full text] [doi: 10.3390/ijerph17062127] [Medline: 32210042]
42. Ostiguy P. Populism: a socio-cultural approach. In: Kaltwasser CR, Taggart P, Espejo PO, Ostiguy P, editors. The Oxford Handbook of Populism. Oxford, UK: Oxford University Press; 2017.
43. Kokorelias KM, Wasilewski MB, Flanagan A, Zhabokritsky A, Singh H, Dove E, et al. Co-creating socio-culturally-appropriate virtual geriatric care for older adults living with HIV: a community-based participatory, intersectional protocol. Int J Qual Methods 2023 Oct 06;22 [FREE Full text] [doi: 10.1177/16094069231205189]
44. Alegria M, Pescosolido BA, Williams S, A G. Culture, race/ethnicity and disparities: fleshing out the socio-cultural framework for health services disparities. In: Pescosolido B, Martin J, McLeod J, Rogers A, editors. Handbook of the Sociology of Health, Illness, and Healing. New York, NY: Springer; 2011.
45. Goldstein MZ, Griswold K. Cultural sensitivity and aging. Psychiatr Serv 1998 Jun;49(6):769-771. [doi: 10.1176/ps.49.6.769] [Medline: 9634155]
46. Xakellis G, Brangman SA, Hinton WL, Jones VY, Masterman D, Pan CX, et al. Curricular framework: core competencies in multicultural geriatric care. J Am Geriatr Soc 2004 Jan 24;52(1):137-142. [doi: 10.1111/j.1532-5415.2004.52024.x] [Medline: 14687329]
47. Periyakoil VS. Building a culturally competent workforce to care for diverse older adults: scope of the problem and potential solutions. J Am Geriatr Soc 2019 May 10;67(S2):S423-S432. [doi: 10.1111/jgs.15939] [Medline: 31074850]
48. Iheneche E. Cultural diversity and cultural competency: new issues for elderly care and services. Arcada University of Applied Science. 2010. URL: https://www.theseus.fi/bitstream/handle/10024/7063/Iheneche_Ernest.pdf?sequence=1 [accessed 2025-05-01]
49. Freshman B, O'Lawrence H, Montion E. The workforce caring for an aging population: addressing the crisis of inadequate geriatric training among healthcare professionals. Workforce Educ Forum 2014;37(1):61-86 [FREE Full text]
50. Sangarlangkarn A, Avihingsanon A, Appelbaum JS. Application of geriatric principles and care models in HIV and aging. Interdiscip Top Gerontol Geriatr 2017;42:119-133. [doi: 10.1159/000448549] [Medline: 27875828]
51. Singh HK, Del Carmen T, Freeman R, Glesby MJ, Siegler EL. From one syndrome to many: incorporating geriatric consultation into HIV care. Clin Infect Dis 2017 Aug 01;65(3):501-506. [doi: 10.1093/cid/cix311] [Medline: 28387803]
52. Montoya ID, Whitsett DD. New frontiers and challenges in HIV research among older minority populations. J Acquir Immune Defic Syndr 2003 Jun 01;33 Suppl 2(Sup 2):S218-S221. [doi: 10.1097/00126334-200306012-00019] [Medline: 12853871]
53. Liboro RM, Bell S, Ranuschio B, Barnes L, Despres J, Sedere A, et al. Barriers and facilitators to promoting resilience to HIV/AIDS: a qualitative study on the lived experiences of HIV-positive, racial and ethnic minority, middle-aged and older men who have sex with men from Ontario, Canada. Int J Environ Res Public Health 2021 Jul 30;18(15):8084 [FREE Full text] [doi: 10.3390/ijerph18158084] [Medline: 34360376]
54. Dubov A, Krakower DS, Rockwood N, Montgomery S, Shoptaw S. Provider implicit bias in prescribing HIV pre-exposure prophylaxis (PrEP) to people who inject drugs. J Gen Intern Med 2023 Oct 24;38(13):2928-2935 [FREE Full text] [doi: 10.1007/s11606-023-08040-7] [Medline: 36964426]
55. Geter A, Herron AR, Sutton MY. HIV-related stigma by healthcare providers in the United States: a systematic review. AIDS Patient Care STDS 2018 Oct;32(10):418-424 [FREE Full text] [doi: 10.1089/apc.2018.0114] [Medline: 30277814]
56. Singleton MC, Green DC, Enguidanos SM. Identifying healthcare stereotype threat in older gay men living with HIV. J Appl Gerontol 2023 Sep 11;42(9):1965-1973. [doi: 10.1177/07334648231167944] [Medline: 37040260]
57. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. BMC Med Ethics 2017 Mar 01;18(1):19 [FREE Full text] [doi: 10.1186/s12910-017-0179-8] [Medline: 28249596]
58. Montoya A. From “the people” to “the human”: HIV/AIDS, neoliberalism, and the economy of virtue in contemporary Vietnam. positions 2012;20(2):561-591. [doi: 10.1215/10679847-1538515]
59. Sankar A, Nevedal A, Neufeld S, Berry R, Luborsky M. What do we know about older adults and HIV? A review of social and behavioral literature. AIDS Care 2011 Oct 14;23(10):1187-1207 [FREE Full text] [doi: 10.1080/09540121.2011.564115] [Medline: 21939401]

60. Baim-Lance A, Angulo M, Chiasson MA, Lekas HM, Schenkel R, Villarreal J, et al. Challenges and opportunities of telehealth digital equity to manage HIV and comorbidities for older persons living with HIV in New York State. BMC Health Serv Res 2022 May 06;22(1):609 [FREE Full text] [doi: 10.1186/s12913-022-08010-5] [Medline: 35524251]
61. Lesko CR, Bengtson AM. HIV and COVID-19: intersecting epidemics with many unknowns. Am J Epidemiol 2021 Jan 04;190(1):10-16 [FREE Full text] [doi: 10.1093/aje/kwaa158] [Medline: 32696057]
62. Flickinger TE, DeBolt C, Xie A, Kosmacki A, Grabowski M, Waldman AL, et al. Addressing stigma through a virtual community for people living with HIV: a mixed methods study of the PositiveLinks mobile health intervention. AIDS Behav 2018 Oct 7;22(10):3395-3406 [FREE Full text] [doi: 10.1007/s10461-018-2174-6] [Medline: 29882048]
63. Rao D, Frey S, Ramaiya M. eHealth for stigma reduction efforts designed to improve engagement in care for people living with HIV. Curr HIV/AIDS Rep 2018 Dec 19;15(6):397-402 [FREE Full text] [doi: 10.1007/s11904-018-0414-z] [Medline: 30232579]
64. Lewis MA, Harshbarger C, Bann C, Burrus O, Peinado S, Garner BR, et al. Positive Health Check evaluation: a type 1 hybrid design randomized trial to decrease HIV viral loads in patients seen in HIV primary care. Contemp Clin Trials 2020 Sep;96:106097. [doi: 10.1016/j.cct.2020.106097] [Medline: 32738408]
65. Armstrong WS, Agwu AL, Barrette EP, Ignacio RB, Chang JJ, Colasanti JA, et al. Innovations in human immunodeficiency virus (HIV) care delivery during the coronavirus disease 2019 (COVID-19) pandemic: policies to strengthen the ending the epidemic initiative-a policy paper of the Infectious Diseases Society of America and the HIV Medicine Association. Clin Infect Dis 2021 Jan 23;72(1):9-14 [FREE Full text] [doi: 10.1093/cid/ciaa1532] [Medline: 33035296]
66. Cohn WF, Canan CE, Knight S, Waldman AL, Dillingham R, Ingersoll K, et al. An implementation strategy to expand mobile health use in HIV care settings: rapid evaluation study using the consolidated framework for implementation research. JMIR Mhealth Uhealth 2021 Apr 28;9(4):e19163 [FREE Full text] [doi: 10.2196/19163] [Medline: 33908893]
67. Walsh JL, Quinn KG, Hirshfield S, John SA, Algiers O, Al-Shalby K, et al. Acceptability, feasibility, and preliminary impact of 4 remotely-delivered interventions for rural older adults living with HIV. AIDS Behav 2024 Apr 03;28(4):1401-1414. [doi: 10.1007/s10461-023-04227-5] [Medline: 38170275]
68. Marent B, Henwood F, Darking M, EmERGE Consortium. Ambivalence in digital health: co-designing an mHealth platform for HIV care. Soc Sci Med 2018 Oct;215:133-141. [doi: 10.1016/j.socscimed.2018.09.003] [Medline: 30232053]
69. Liang C, Hutson JA, Keyes O. Surveillance, stigma and sociotechnical design for HIV. First Monday 2020 Oct 5;25(10). [doi: 10.5210/fm.v25i10.10274]
70. Massaroni V, Delle Donne V, Ciccarelli N, Ciccullo A, Borghetti A, Faliero D, et al. Use of telehealth for HIV care in Italy: are doctors and patients on the same page? A cross-sectional study. Int J Med Inform 2021 Dec;156:104616. [doi: 10.1016/j.ijmedinf.2021.104616] [Medline: 34695728]
71. Kim JY, Barth SK, Monroe AK, Ahsan S, Kovacic J, Senn S, et al. The impact of COVID-19 on the HIV continuum of care: challenges, innovations, and opportunities. Expert Rev Anti Infect Ther 2023 Jul 27;21(8):831-846. [doi: 10.1080/14787210.2023.2239503] [Medline: 37470436]
72. Phan JM, Kim S, Linh Đ, Cosimi LA, Pollack TM. Telehealth interventions for HIV in low- and middle-income countries. Curr HIV/AIDS Rep 2022 Dec 26;19(6):600-609 [FREE Full text] [doi: 10.1007/s11904-022-00630-0] [Medline: 36156183]
73. Shaw J, Brewer LC, Veinot T. Recommendations for health equity and virtual care arising from the COVID-19 pandemic: narrative review. JMIR Form Res 2021 Apr 05;5(4):e23233 [FREE Full text] [doi: 10.2196/23233] [Medline: 33739931]
74. Eyrich NW, Andino JJ, Fessell DP. Bridging the digital divide to avoid leaving the most vulnerable behind. JAMA Surg 2021 Aug 01;156(8):703-704. [doi: 10.1001/jamasurg.2021.1143] [Medline: 33769441]
75. Luque AE, Corales R, Fowler RJ, DiMarco J, van Keken A, Winters P, et al. Bridging the digital divide in HIV care: a pilot study of an iPod personal health record. J Int Assoc Provid AIDS Care 2013 Sep 10;12(2):117-121 [FREE Full text] [doi: 10.1177/1545109712457712] [Medline: 22965693]
76. Choukou MA, Sanchez-Ramirez DC, Pol M, Uddin M, Monnin C, Syed-Abdul S. COVID-19 infodemic and digital health literacy in vulnerable populations: a scoping review. Digit Health 2022 Feb 10;8:20552076221076927 [FREE Full text] [doi: 10.1177/20552076221076927] [Medline: 35223076]
77. Nixon E, Cooper V, Donetto S, Youssef E, Robert G. Co-designing health services for people living with HIV who have multimorbidity: a feasibility study. Br J Nurs 2023 Jun 08;32(11):S15-S21. [doi: 10.12968/bjon.2023.32.11.s15]
78. Yang C, Teh YE, Chua NG, Lee KL, Ng RQ. An overview of multimorbidity and polypharmacy in older people living with HIV. Geriatr Gerontol Int 2024 Mar 08;24 Suppl 1(S1):49-59. [doi: 10.1111/ggi.14717] [Medline: 37940135]
79. Kokorelias KM, Valentine D, Dove EM, Brown P, McKinlay S, Sheppard CL, et al. Exploring the perspectives of older adults living with HIV on virtual care: qualitative study. JMIR Aging 2024 Dec 04;7:e65730 [FREE Full text] [doi: 10.2196/65730] [Medline: 39630418]
80. Zamenopoulos T, Alexiou K. Co-Design as Collaborative Research. Bristol, England: University of Bristol and the AHRC Connected Communities Programme; Jan 09, 2019:1-53.
81. Castro EM, Malfait S, van Regenmortel T, van Hecke A, Sermeus W, Vanhaecht K. Co-design for implementing patient participation in hospital services: a discussion paper. Patient Educ Couns 2018 Jul;101(7):1302-1305. [doi: 10.1016/j.pec.2018.03.019] [Medline: 29602511]

82. Maye L, Claisse C. Co-design within and between communities in cultural heritage: current and open questions. Multimodal Technol Interact 2022 Dec 20;7(1):1. [doi: 10.3390/mti7010001]
83. Armoon B, Higgs P, Fleury MJ, Bayat AH, Moghaddam LF, Bayani A, et al. Socio-demographic, clinical and service use determinants associated with HIV related stigma among people living with HIV/AIDS: a systematic review and meta-analysis. BMC Health Serv Res 2021 Sep 22;21(1):1004 [FREE Full text] [doi: 10.1186/s12913-021-06980-6] [Medline: 34551772]
84. Donetto S, Tsianakas V, Robert G. Using experience-based co-design (EBCD) to improve the quality of healthcare: mapping where we are now and establishing future directions. King's College London. 2014. URL: https://www.scirp.org/reference/referencespapers?referenceid=3581170 [accessed 2025-05-01]
85. Raynor DK, Ismail H, Blenkinsopp A, Fylan B, Armitage G, Silcock J. Experience-based co-design-Adapting the method for a researcher-initiated study in a multi-site setting. Health Expect 2020 Jun 11;23(3):562-570 [FREE Full text] [doi: 10.1111/hex.13028] [Medline: 32045087]
86. Howrey BT, Thompson BL, Borkan J, Kennedy LB, Hughes LS, Johnson BH, et al. Partnering with patients, families, and communities. Fam Med 2015 Sep;47(8):604-611 [FREE Full text] [Medline: 26382118]
87. Lindamer LA, Lebowitz B, Hough RL, Garcia P, Aguirre A, Halpain MC, et al. Establishing an implementation network: lessons learned from community-based participatory research. Implement Sci 2009 Mar 31;4(1):17 [FREE Full text] [doi: 10.1186/1748-5908-4-17] [Medline: 19335915]
88. Eaton AD. Filmed simulation to train peer researchers in community-based participatory research. Soc Work Res 2019 Sep;43(3):195-199. [doi: 10.1093/swr/svz011]
89. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
90. Brown P, Singh H, Su E, Sirisegaram L, Munce SE, Eaton AD, et al. Understanding the use of co-design methods for research involving older adults living with HIV: a scoping review protocol. PLoS One 2024 May 30;19(5):e0303580 [FREE Full text] [doi: 10.1371/journal.pone.0303580] [Medline: 38814951]
91. Robert G. Participatory action research: using experience-based co-design to improve the quality of healthcare services. In: Ziebland S, Calabrase J, Coulter A, Locock L, editors. Understanding and Using Experiences of Health and Illness. Oxford, UK: Oxford University Press; 2013:138-149.
92. Pearce V, Baraitser P, Smith G, Greenhalgh T. Experience-based co-design. In: Greenhalgh T, Humphrey C, Woodard F, editors. User Involvement in Health Care. Hoboken, NJ: Blackwell Publishing Ltd; 2010.
93. Marshall PA, Rotimi C. Ethical challenges in community-based research. Am J Med Sci 2001 Nov;322(5):241-245. [doi: 10.1097/00000441-200111000-00001]
94. Moser CA. Quota sampling. J R Stat Soc Ser A 1952;115(3):411-423. [doi: 10.2307/2980740]
95. Suri H. Purposeful sampling in qualitative research synthesis. Qual Res J 2011;11(2):63-75. [doi: 10.3316/qrj1102063]
96. Calabrese R, Cohen E, Miller D. Creating a healthy workplace culture using an appreciative inquiry 4-D cycle. Organ Manag J 2013 Sep;10(3):196-207. [doi: 10.1080/15416518.2013.831706]
97. Faure M, Rosenzweig J, van Tiem D. Appreciative inquiry. In: Watkins R, Leigh D, editors. Handbook of Improving Performance in the Workplace: Selecting and Implementing Performance Interventions. Hoboken, NJ: John Wiley & Sons; 2009.
98. Kane M, Trochim WM. Concept Mapping for Planning and Evaluation. Thousand Oaks, CA: SAGE Publications; 2007.
99. Löhr K, Weinhardt M, Sieber S. The "world café" as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Apr 24;19. [doi: 10.1177/1609406920916976]
100. Anderson R. Thematic content analysis (TCA): descriptive presentation of qualitative data. rosemarieanderson.com. 2007. URL: https://rosemarieanderson.com/wp-content/uploads/2014/08/ThematicContentAnalysis.pdf [accessed 2025-05-01]
101. Nicmanis M. Reflexive content analysis: an approach to qualitative data analysis, reduction, and description. Int J Qual Methods 2024;23. [doi: 10.1177/16094069241236603]
102. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
103. DeCuir-Gunby JT, Marshall PL, McCulloch AW. Developing and using a codebook for the analysis of interview data: an example from a professional development research project. Field Methods 2010 Dec 27;23(2):136-155. [doi: 10.1177/1525822x10388468]
104. Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: development and refinement of a codebook. Comp Psychol 2014 Feb 20;3. [doi: 10.2466/03.cp.3.4]
105. Viswanathan M, Ammerman A, Eng E, Garlehner G, Lohr KN, Griffith D, et al. Community-based participatory research: assessing the evidence. Evid Rep Technol Assess (Summ) 2004 Aug(99):1-8. [Medline: 15460504]
106. Natow RS. The use of triangulation in qualitative studies employing elite interviews. Qual Res 2019 Feb 19;20(2):160-173. [doi: 10.1177/1468794119830077]
107. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Aug 26;41(5):545-547. [doi: 10.1188/14.onf.545-547]

108. Cutcliffe JR, McKenna HP. Establishing the credibility of qualitative research findings: the plot thickens. J Adv Nurs 1999 Aug 25;30(2):374-380. [doi: 10.1046/j.1365-2648.1999.01090.x] [Medline: 10457239]
109. Passmore SR, Kisicki A, Gilmore-Bykovskyi A, Green-Harris G, Edwards DF. "There's not much we can do…" researcher-level barriers to the inclusion of underrepresented participants in translational research. J Clin Transl Sci 2022 Nov 11;6(1):e4 [FREE Full text] [doi: 10.1017/cts.2021.876] [Medline: 35154814]
110. Rush KL, Seaton CL, Corman K, Hawe N, Li EP, Dow-Fleisner SJ, et al. Virtual care prior to and during COVID-19: cross-sectional survey of rural and urban adults. JMIR Form Res 2022 Aug 22;6(8):e37059 [FREE Full text] [doi: 10.2196/37059] [Medline: 35849794]
111. Perri M, Schmidt RA, Guta A, Kaminski N, Rudzinski K, Strike C. COVID-19 and the opportunity for gender-responsive virtual and remote substance use treatment and harm reduction services. Int J Drug Policy 2022 Oct;108:103815 [FREE Full text] [doi: 10.1016/j.drugpo.2022.103815] [Medline: 35969920]
112. Griffin B. Virtual care in Canada: snapshots of innovative virtual care. Digital Health Canada. 2019 Nov. URL: https://divisionsbc.ca/sites/default/files/Divisions/Kootenay%20Boundary/Misc%20-%20events-imags/Digital%20Health%20Canada%20Virtual%20Care%20in%20Canada%20vFinal%20DEC2-2019%5B1%5D.pdf [accessed 2025-05-01]
113. Gari S, Doig-Acuña C, Smail T, Malungo JR, Martin-Hilber A, Merten S. Access to HIV/AIDS care: a systematic review of socio-cultural determinants in low and high income countries. BMC Health Serv Res 2013 May 28;13(1):198 [FREE Full text] [doi: 10.1186/1472-6963-13-198] [Medline: 23714167]
114. Rose MA. Planning HIV education programs for older adults: cultural implications. J Gerontol Nurs 2004 Mar;30(3):34-39. [doi: 10.3928/0098-9134-20040301-11] [Medline: 15061452]
115. Green T, Bonner A, Teleni L, Bradford N, Purtell L, Douglas C, et al. Use and reporting of experience-based codesign studies in the healthcare setting: a systematic review. BMJ Qual Saf 2020 Jan 23;29(1):64-76. [doi: 10.1136/bmjqs-2019-009570] [Medline: 31548278]
116. Webber R, Partridge R, Grindell C. The creative co-design of low back pain education resources. Evid Policy 2022;18(2):436-453 [FREE Full text] [doi: 10.1332/174426421x16437342906266]
117. Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based co-design and healthcare improvement: realizing participatory design in the public sector. Des Journal 2015 May 07;18(2):227-248. [doi: 10.2752/175630615X14212498964312]
118. Mitzner TL, Fausset CB, Boron JB, Adams AE, Dijkstra K, Lee CC, et al. Older adults' training preferences for learning to use technology. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep 01;52(26):2047-2051 [FREE Full text] [doi: 10.1177/154193120805202603] [Medline: 25309139]
119. Hickman J, Rogers WA, Fisk AD. Training older adults to use new technology. J Gerontol B Psychol Sci Soc Sci 2007 Jun;62 Spec No 1(Special_Issue_1):77-84. [doi: 10.1093/geronb/62.special_issue_1.77] [Medline: 17565168]
120. Patterson PB, Roddick J, Pollack CA, Dutton DJ. Virtual care and the influence of a pandemic: necessary policy shifts to drive digital innovation in healthcare. Healthc Manage Forum 2022 Sep 30;35(5):272-278 [FREE Full text] [doi: 10.1177/08404704221110084] [Medline: 35775144]
121. Panaite AC, Desroches OA, Warren É, Rouly G, Castonguay G, Boivin A. Engaging with peers to integrate community care: knowledge synthesis and conceptual map. Health Expect 2024 Apr 03;27(2):e14034 [FREE Full text] [doi: 10.1111/hex.14034] [Medline: 38567865]

Abbreviations

CBPR: community-based participatory research
EBCD: experience-based co-design

Edited by R Bjarnadottir; submitted 02.10.24; peer-reviewed by AL Nguyen, A Peeples; comments to author 10.03.25; revised version received 17.03.25; accepted 21.04.25; published 27.05.25.

Please cite as:
Kokorelias KM, Wasilewski MB, Valentine D, Eaton AD, Dove E, Brown P, McKinlay S, Sheppard CL, Su E, Singh HK, Flanagan A, Zhabokritsky A, Abdelhalim R, Parpia R, Zewude R, Jamieson L, Walmsley S, Sirisegaram L
Socioculturally Appropriate Internet-Based Geriatric Care Model for Older Adults Living With HIV: Experience-Based Co-Design Approach
JMIR Aging 2025;8:e67122
URL: https://aging.jmir.org/2025/1/e67122
doi:10.2196/67122
PMID:

Original Paper

Longitudinal Remote Sleep and Cognitive Research in Older Adults With Mild Cognitive Impairment and Dementia: Prospective Feasibility Cohort Study

Victoria Grace Gabb[1,2,3], BSc, MSc; Jonathan Blackman[1,2,3], BSc, BMBS, MRCP, MRCPsych, PhD; Hamish Morrison[1,2], BMedSci, MBChB, MRCP; Haoxuan Li[1,2,4,5,6], BSc, MBBS; Adrian Kendrick[7], BA, PgD, PhD; Nicholas Turner[8], BSc, MSc, PhD; Rosemary Greenwood[9,10], BSc, MSc; Bijetri Biswas[11], BSc, MSc; Amanda Heslegrave[12,13], BSc, PhD; Elizabeth Coulthard[1,2,3], BA, MBBS, FRCP, PhD

[1]Bristol Medical School, Translational Health Sciences, University of Bristol, Bristol, United Kingdom
[2]Neurology Department, Southmead Hospital, North Bristol NHS Trust, Bristol, United Kingdom
[3]Mental Health, NIHR Bristol Biomedical Research Centre, Bristol, United Kingdom
[4]Bristol Royal Infirmary, University Hospitals Bristol and Weston NHS Trust, Bristol, United Kingdom
[5]King's College London, London, United Kingdom
[6]King's College Hospital, London, United Kingdom
[7]University of the West of England, Bristol, United Kingdom
[8]Bristol Trials Centre, Bristol Medical School, University of Bristol, Bristol, United Kingdom
[9]Research & Innovation, University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom
[10]York Trials Unit, Department of Health Sciences, University of York, York, United Kingdom
[11]School of Electrical, Electronic and Mechanical Engineering, University of Bristol, Bristol, United Kingdom
[12]Institute of Neurology, University College London, London, United Kingdom
[13]UK Dementia Research Institute, University College London, London, United Kingdom

Corresponding Author:
Victoria Grace Gabb, BSc, MSc
Bristol Medical School
Translational Health Sciences
University of Bristol
Bristol Brain Centre
Elgar House
Bristol, BS10 5NB
United Kingdom
Phone: 44 01174560700
Email: victoria.gabb@bristol.ac.uk

Abstract

Background: Sleep holds promise as a modifiable risk factor for neurodegenerative diseases and dementia. Clinical trials to modify sleep in people at risk of or in the early stages of dementia are needed. Monitoring natural sleep from home could support pragmatic and decentralized large-scale clinical trials. However, whether longitudinal sleep research can be successfully delivered remotely in this population has not been established yet.

Objective: We investigated the feasibility of remote longitudinal research using wearable devices, web-based cognitive tasks, and a smartphone app to record sleep and cognition in older adults with mild cognitive impairment (MCI) or dementia.

Methods: Older adults with MCI or dementia due to Alzheimer disease or Lewy body disease and cognitively healthy participants completed at-home sleep and circadian monitoring (digital sleep diaries, actigraphy, wearable sleep electroencephalography, and saliva samples) and digital cognitive assessments for 8 weeks. Feasibility outcomes included recruitment, retention, and data completeness.

Results: In total, 41 participants consented (n=10, 24% participants with Alzheimer disease; n=11, 27% participants with Lewy body disease; and n=20, 49% controls). There were predominantly male and White British participants, with a mean age of 70.9

(SD 5.9) years. Retention was very high, with 40 (98%) participants completing 8 weeks of remote monitoring. Data completeness for sleep electroencephalography was 91% and ranged from 79% to 97% for all remote tasks and was overall high across all participant subgroups. In total, 30% (12/40) of participants reported receiving external support with completing study tasks.

Conclusions: High rates of retention, data completeness, and data quality suggested that longitudinal multimodal sleep and cognitive profiling using novel and remote monitoring technology is feasible in older adults with MCI and dementia and healthy older adults, even without study partner support. Remote monitoring should be considered for mechanistic and interventional trials. Careful consideration should be given to how to ensure remote monitoring technologies reduce burden and enhance inclusivity, particularly in communities traditionally underserved by research and those with lower digital literacy.

International Registered Report Identifier (IRRID): RR2-10.2196/52652

(JMIR Aging 2025;8:e72824) doi:10.2196/72824

KEYWORDS

feasibility; sleep; remote study design; mild cognitive impairment; dementia; electroencephalography; actigraphy; saliva; digital biomarkers

Introduction

Background

Sleep disturbances such as insomnia, fragmented sleep, daytime sleepiness, and sleep-disordered breathing are common features of Alzheimer disease (AD) and Lewy body disease (LBD) and often appear early in the disease course and before clinical diagnosis [1,2]. Short and fragmented sleep are associated with processes linked to neurodegeneration, including reduced glymphatic clearance of waste [3], increased amyloid beta (Aβ) and tau burden [4], increased neuroinflammation [5], and impaired cardiovascular health [6]. Several small studies of sleep apnea treatment have demonstrated improvement in cognitive outcomes and blood biomarkers of Aβ and tau, suggesting that sleep interventions could improve prognosis in mild cognitive impairment (MCI) and dementia [7-9]. Poor quality or insufficient sleep in midlife increases the risk of all-cause dementia and MCI [10,11], indicating that improving sleep may also protect against dementia. Large-scale clinical studies are needed to enhance our mechanistic understanding of sleep, confirm the most promising therapeutic targets, and monitor the effectiveness and safety profile of sleep interventions [12,13].

However, selecting sleep assessment tools is challenging, particularly in populations with cognitive impairment. Self-report is convenient, inexpensive, and scalable, and has often been used to examine sleep in people with dementia and MCI [14]; however, it is potentially prone to recall bias due to memory deficits or anosognosia. Self-report also correlates poorly with objective sleep measures, particularly in participants with MCI or dementia [15,16] and those with subjectively poor sleep [17], and cannot inform on key components of sleep such as sleep staging or microarchitecture. Polysomnography (PSG) is typically considered the gold standard for sleep measurement, as it provides rich objective sleep data. However, PSG typically requires expert setup, analysis, and a controlled clinic environment, meaning it is expensive, not easily scalable, and therefore typically used for 1 or a few nights [18,19]. PSG setup, especially in an artificial environment, may also not reflect natural sleep [19] as it disrupts usual routines. Longitudinal data collection in sleep research would be beneficial for monitoring clinical trials and disease progression and could account for variation in sleep from external factors such as acute illness or stress, as well as natural intraindividual sleep variation [19].

Wearable devices, smartphone apps, and telemedicine, apart from actigraphy devices, have rarely been used in MCI or dementia research [20], but offer an opportunity to collect objective and subjective data longitudinally in a natural setting [21], often at relatively low cost. Detailed and accurate sleep analysis can now be achieved through wireless technology, including electroencephalography (EEG) headbands and overnight pulse oximetry [21,22]. While actigraphy has been used in dementia research, EEG headbands and pulse oximetry are less well tested, especially in the earlier stages of impairment [23-25]. There is also increasing interest in remote cognitive testing and digital biomarkers for diagnosis and monitoring progression [26]. Despite this, the adoption of digital health technologies into neurology clinical trials, particularly in older people, has been slow [27].

Remote sleep and cognitive data collection has the potential to decentralize clinical trials, making research more convenient for participants and reducing the costs, participant burden, and carbon footprint associated with study visits, while enabling real-time longitudinal data collection of treatment effects. Improved digital access and literacy among older adults [28] and the use of technology among patients living with MCI or dementia to support independent living and for recreation indicate increasing acceptance of technology [29]. However, not all older adults are comfortable using technology, and changes in cognition, sensory processing, and communication might negatively impact the usability and acceptability of novel devices for research purposes in people with cognitive impairment [29,30]. The few studies that have tested wearable devices and digital health technologies for sleep and dementia research in the home have predominantly collected feasibility data for a single device across only a few nights [31] and required support from a study partner or care home staff [32-34]. Before trials invest in digital health technologies and remote study designs, it is important to know whether participants can and are willing to engage and provide data in such studies over a longer period and how much support might be required.

Objectives

This study aimed to establish the feasibility of predominantly technology-based longitudinal sleep and cognitive assessments from home in older adults with and without MCI or dementia.

Methods

Overview

The Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study was a prospective longitudinal observational cohort study involving remote sleep and cognitive monitoring of participants with MCI or dementia and age-matched cognitively healthy controls. The full study protocol has been previously published [35]. The study has been reported in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [36] (Multimedia Appendix 1).

Ethical Considerations

This study was approved by the Health Research Authority (Yorkshire and the Humber-Bradford Leeds Research Ethics Committee, reference 21/YH/0177). This study was conducted in accordance with the principles of the Declaration of Helsinki. All participants provided written informed consent before engaging in any study activities and were reminded of their right to withdrawal without giving a reason and that withdrawal would not affect their health care. Following consent, data were pseudonymized with participants allocated a participant identifier code to maintain privacy. Participants were not compensated for their involvement in the study; however, travel expenses for research visits outside of usual clinic visits were reimbursed. All researchers engaging with participants were trained in good clinical practice.

Study Population

The RESTED study recruited community-dwelling adults aged ≥50 years with internet access at home. Participants were recruited to 1 of 3 participant subgroups according to clinical diagnosis meeting standardized diagnostic criteria [37-40]. Participants with MCI or mild dementia due to probable AD were recruited to the AD group, participants with MCI or mild dementia due to probable LBD or Parkinson disease (PD) were recruited to the LBD group, and sex- and age-matched individuals with no known neurodegenerative conditions or cognitive impairment were recruited as controls. Exclusion criteria included advanced dementia, acute or terminal illness, and significant unrelated comorbidities that might interfere with sleep, except for obstructive sleep apnea (OSA) if participants were already undergoing treatment.

Participants were recruited from the city of Bristol, United Kingdom, and the surrounding areas via cognitive and movement disorders clinics and research volunteer databases, including Join Dementia Research. Our original target sample size was 75 participants [41]. Due to the COVID-19 pandemic, the study was delayed, and the budget was partially reallocated to studies that required no patient contact [12,13,42]. Therefore, we revised the target sample size to 40 participants and recruited from a single site (North Bristol NHS Trust).

We did not recruit or require participants to have a study partner. Participants were welcome to invite someone to support them, and study support was recorded.

Study Procedures

Screening

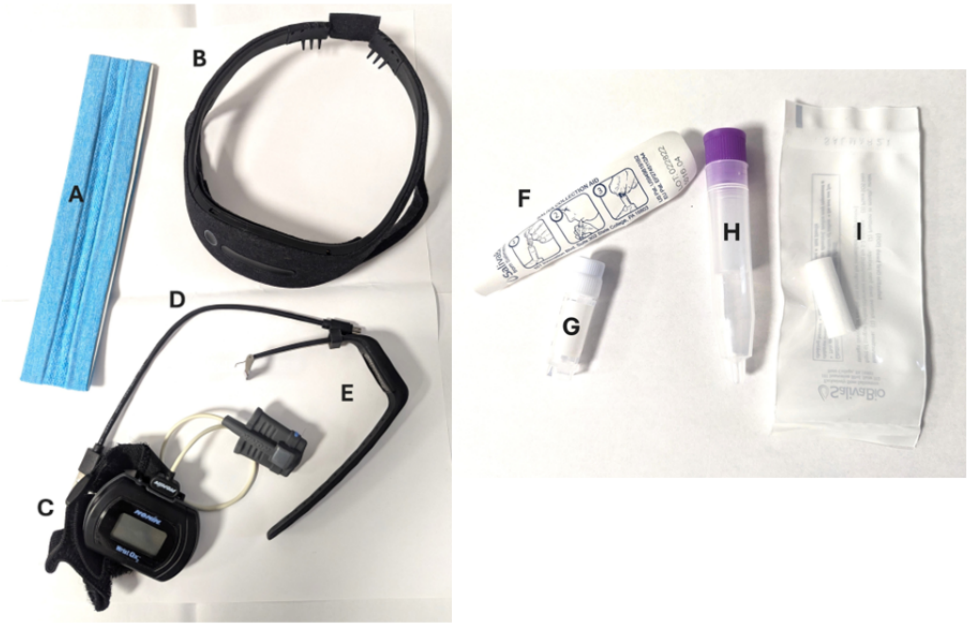
Participants were prescreened for eligibility over the telephone and invited to complete consent and screening in person. Participants who scored less than 11 out of 30 on the Montreal Cognitive Assessment (MoCA) [43] at screening were considered too clinically impaired to participate and were withdrawn.

Baseline Assessments

Eligible participants completed baseline questionnaires with a researcher to assess sleep quality (Pittsburgh Sleep Quality Index [PSQI]) [44]; daytime sleepiness (Epworth Sleepiness Scale) [45]; OSA risk (STOP-Bang) [46]; symptoms of depression (Geriatric Depression Scale-15 item) [47]; anxiety (Generalized Anxiety Disorder-7 item) [48]; and apathy (Apathy Evaluation Scale-Self) [49]. Demographic information and medical histories were also recorded.

Participants were provided with a study kit (Figure 1) consisting of an Axivity AX3 actigraphy watch (Axivity Ltd); Dreem 2 wireless sleep EEG headband (Dreem) [23]; a USB charger; a home saliva collection kit for passive drool (for a dim-light melatonin onset assay); oral swabs (for a cortisol awakening response assay); and an overnight pulse oximeter. Participants who requested a study device were provided with a tablet. Details on devices and the study kit are provided in Multimedia Appendix 2. All participants were provided with a printed participant guide with instructions for each remote study task and research team contact details.

Figure 1. The study kit provided to the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) participants. Participants were additionally required to use a smartphone or tablet to complete digital assessments. A: Sweatband to ensure tight fitting of Dreem 2; B: Dreem 2 electroencephalography (EEG) headband; C: Nonin 3150 WristOx 2 oximeter; D: USB-C charging cable (for Dreem 2 and AX3); E: Axivity AX3 actigraph and wrist strap; F: Saliva collection aid for passive drool; G: Collection tube for passive drool; H: Storage tube for oral swab; I: Oral swab.



At baseline, participants were provided with support to download and use the MyDignio (Dignio AS) app, a telemedicine cloud-based software tool that was used to deliver digital sleep diaries, questionnaires, and reminders to complete study tasks. Participants were also familiarized with the study schedule and tasks and offered support and training ad hoc throughout the study.

Main Study Period

Participants completed 56 days of continuous sleep and regular cognitive monitoring in their own home using an Axivity AX3 actigraphy watch and daily sleep diaries [50] delivered via the MyDignio app. Participants were scheduled to complete a set of 3 web-based cognitive tests (choice reaction time, forward digit span, and self-ordered search) twice weekly on a bespoke study version of the web-based assessment platform Cognitron [51].

For 7 days during the 8-week period, participants also completed an "intensive week," consisting of daily cognitive tasks on Cognitron and nighttime sleep recordings using Dreem 2, a wireless sleep EEG headband. The Dreem sleep staging algorithm has comparable accuracy to manual sleep expert scoring of PSG [23], including in cognitively healthy older adults as well as patients with AD [52] and PD [53]. Sleep recordings were initiated by the participant at their natural bedtime each night and terminated after natural awakening each morning. Data were uploaded via Bluetooth and Wi-Fi to a server accessible to the research team. Additional study tasks during the intensive week included 4 verbal memory assessments involving immediate and delayed free recall and a target-distractor recognition task with a researcher via videoconferencing software and serial saliva samples across 1 evening to assess dim-light melatonin onset and 1 morning to assess cortisol awakening response.

Participants were also invited to undergo 2 nights of pulse oximetry for sleep apnea screening and a blood test for plasma biomarker analysis of Aβ42:40, phosphorylated-tau 181 and 217, neurofilament light chain, and glial fibrillary acidic protein. A protocol amendment approved partway through the study introduced bespoke questionnaires probing study expectations, reasons for participation, experience with technology, and how acceptable they found the intensive week study tasks. Participants were also invited to attend a remote end-of-study interview to share their experiences and asked to return for a 6-month follow-up to complete an MoCA and any outstanding study tasks (eg, missed blood test). Participant feedback from qualitative interviews and questionnaires as well as sleep characteristics for the cohort will be analyzed and presented separately.

Feasibility and Acceptability Outcomes

Recruitment and Retention

Consent, recruitment, and retention rates were recorded and are summarized in a flowchart, alongside reasons for nonparticipation at each stage of the recruitment process.

Data Quality and Completeness

For each remote study task, data completeness was assessed by the average number and percentage of completed tasks or nights' use per participant per participant group (data completeness rate). The number and percentage of individuals who completed the maximum number of data points for a given study task (eg, completed all requested 7 nights of EEG) is also provided. For

EEG, signal quality is reported based on Dreem's automated algorithm. Optimal signal quality is considered ≥85% and good quality is considered ≥70%.

Associations Between Participant Characteristics and Adherence

To explore potential associations between key clinical and demographic variables and adherence, we examined the correlation between 2 continuous measures of adherence (number of sleep diaries completed and mean record quality of Dreem EEG data) and 4 continuous variables which might impact digital literacy or engagement (age at consent, apathy, subjective sleep quality as assessed by the PSQI total score, and baseline cognitive impairment as assessed by MoCA total score).

Study Support and Resource Use

The number of participants who attended study visits with someone to support them, reported having a partner or relative who could support study tasks, and reported receiving support on at least 1 study task were recorded and are summarized descriptively.

Data Analysis

Unless otherwise stated, descriptive statistics are provided as mean (SD) for continuous variables, and frequency and percentage for categorical variables, and provided for the full cohort as well as by participant subgroup (AD vs LBD vs controls). Data analysis was performed using R (version 4.3.1; R Foundation for Statistical Computing) and R Studio software (version 2023.6.0.421; R Foundation for Statistical Computing). Actigraphy data were processed using the open-source AX3/AX6 OMGUI application and analyzed in R using the *GGIR* package [54]. Where available, sleep timing was adjusted using participants' sleep diaries; otherwise, *GGIR* uses the heuristic algorithm looking at distribution of change in z-angle [55] to estimate sleep timing. All files were also visually inspected to check for the accuracy of sleep timing. The quality of the Dreem 2 recordings was assessed by inspecting Dreem's automated record quality index in the sleep report for each night, which indicates the percentage of the recording that is of scorable quality for sleep analysis. Associations between participant characteristics and adherence were calculated using Spearman rank correlation with α set to $P<.05$. Reasons for nonparticipation and missing data are provided where known and mapped to the Capability, Opportunity, Motivation model of behavior change [56]. Where additional data were collected on any outcome (eg, participants completed an additional night of EEG than instructed), additional data points were removed before analysis to avoid biasing feasibility metrics.

Patient and Public Involvement

Patient and public involvement (PPI) was sought from individuals with lived experience of MCI and dementia before and throughout the study. PPI contributors reviewed and improved study documents and advised on the acceptability of adding blood biomarker testing. The PPI group strongly endorsed our additional recruitment materials (including an advertisement poster and a plain English participant information sheet) that we introduced partway through the study, after we received feedback from a prospective participant that standard participant information sheets are too long and use inaccessible language for people with cognitive impairment. We also introduced a reduced study protocol, which involved use of a paper (rather than digital) sleep diary, actigraphy, and EEG, to encourage recruitment of participants who may be more comfortable with less frequent use of technology; however, nobody chose the reduced protocol, they either declined altogether or participated in the full study.

Results

Participant Characteristics and Enrollment

Recruitment

Participant flow through the study is shown in Figure 2. Recruitment was open for 17 months from February 2022 to July 2023, with an average recruitment rate of 3 participants per month. Of 129 individuals identified as potentially eligible before screening, 44 individuals consented to take part, giving a consent rate of 34%. Participants described different motivations for taking part, including wanting to support dementia research, perceiving the study as helpful for themselves or others, and because the study sounded interesting or novel. Reasons for declining to take part in the study are described in Figure 3 and mapped to the Capability, Opportunity, Motivation model of behavior change [56].

Figure 2. Participant flow through the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study. MCI: mild cognitive impairment; MoCA: Montreal Cognitive Assessment.

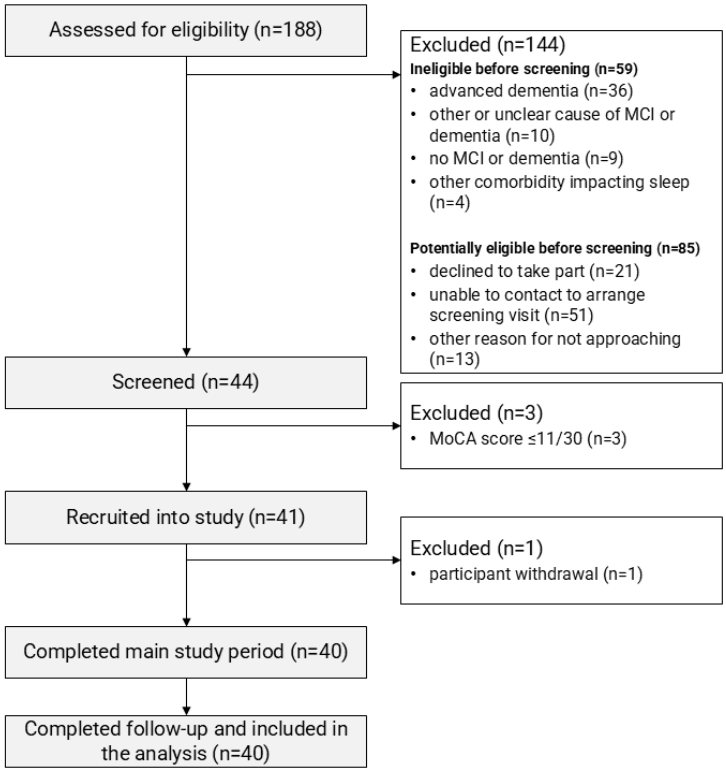
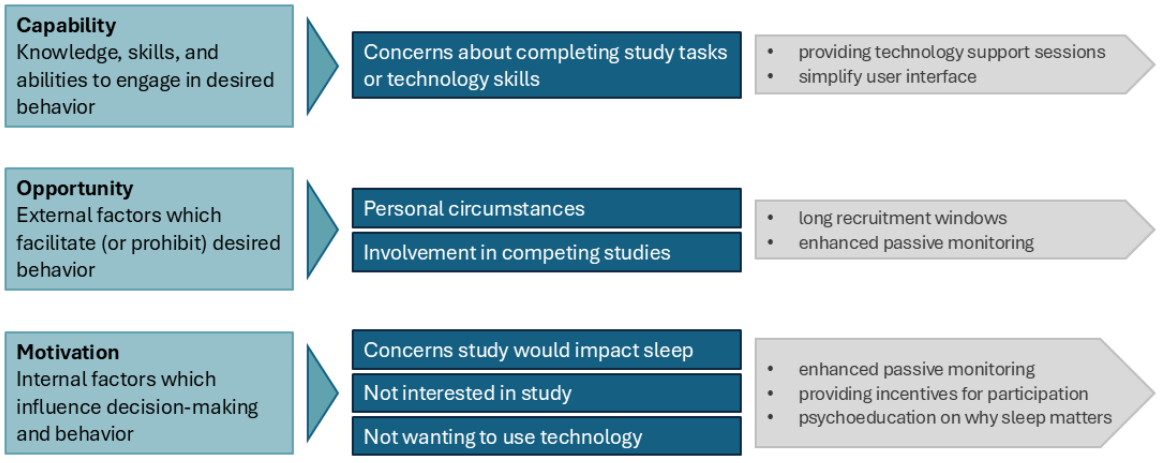


Figure 3. Reasons for declining to participate in the study could be categorized according to the capability, opportunity motivation, behavior (COM-B) model of behavior change in blue. Opportunities to increase capability, opportunity, and motivation to take part in the research are outlined in gray.



Retention

Of 44 individuals who completed screening, 3 scored less than 11 on the MoCA and were withdrawn. 41 participants were eligible and completed all baseline assessments and training in the at-home study tasks. Of the eligible participants, 1 withdrew before starting the at-home study tasks due to personal circumstances and perceived study burden. All remaining 40 participants completed the main 8-week study period and 6-month follow-up and were included in the analysis, giving a retention rate of 98%.

Participant Characteristics

Participant demographics and baseline variables are presented in Table 1. The sample was predominantly made up of male participants with a mean age at consent of 70.9 (SD 5.9, range 57-81) years. Most participants (33/40, 83%) were retired and all identified as White British. In total, 63% (25/40) of participants completed a questionnaire on experiences of digital technology and expectations of the study, which was introduced via an amendment partway through the recruitment period. Of these, most participants reported frequent use of smartphones or other smart technology (18/25, 72%), however only 12% (3/25) of participants had previously used a wearable to monitor their sleep.

Table 1. Participant characteristics for the study cohort.

Characteristics	AD[a] group (n=10)	LBD[b] group (n=10)	Control group (n=20)
Diagnosis at consent, n (%)			
Mild cognitive impairment	4 (40)	6 (60)	0 (0)
Dementia	6 (60)	4 (40)	0 (0)
Blood biomarkers of neurodegeneration or AD pathology (pg/mL), mean (SD)			
Aβ[c]42:40	0.064 (0.006)	0.062 (0.006)	0.072 (0.007)
Phosphorylated-tau 181	26.2 (7.1)	24.9 (12.7)	24.9 (12.7)
Phosphorylated-tau 217	0.7 (0.4)	0.8 (0.4)	0.6 (0.4)
GFAP[d]	102.2 (28.1)	140.9 (56.4)	103.5 (48.1)
NFL[e]	23.9 (18.0)	18.7 (6.5)	13.4 (5.1)
Age at consent (y), mean (SD)	69.2 (7.9)	73.9 (2.8)	70.3 (5.7)
Female participants, n (%)	2 (20)	2 (20)	5 (25)
Employment status, n (%)			
Full-time employment	2 (20)	0 (0)	2 (10)
Part-time employment	0 (0)	0 (0)	3 (15)
Retired	8 (80)	10 (100)	15 (75)
Education, n (%)			
Secondary (≤12 y)	4 (40)	5 (50)	4 (20)
Further or higher (>12 y)	6 (60)	5 (50)	16 (80)
Comorbid diagnoses, n (%)			
Obstructive sleep apnea[f]	0 (0)	1 (10)	2 (10)
Musculoskeletal or pain	3 (30)	3 (30)	8 (40)
Urinary or prostate	3 (30)	2 (20)	3 (15)
Depression or anxiety	1 (10)	0 (0)	1 (5)
Medications, n (%)			
Cognitive enhancer	6 (60)	6 (60)	0 (0)
Dopaminergic	0 (0)	8 (80)	0 (0)
Melatonin	0 (0)	3 (30)	0 (0)
Hypnotic or sedative	0 (0)	2 (20)	1 (5)
Antidepressant	4 (40)	4 (40)	4 (20)
Baseline cognition: MoCA[g]; mean (SD)	23.0 (4.9)	22.0 (4.3)	27.1 (1.4)
Subjective sleep quality: PSQI[h]; mean (SD)	4.3 (2)	7.8 (3)	6.0 (3.9)
Self-reported anxiety: GAD-7[i]; mean (SD)	5.6 (5.2)	3.5 (2.1)	3.6 (4.3)
Self-reported depression: GDS-15[j]; mean (SD)	4.5 (3.3)	4.1 (2.3)	2.2 (2.2)
Self-reported apathy: AES-S[k,l]; mean (SD)	62.2 (6.0)	53.8 (9.9)	62.8 (8.6)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body disease.

[c]Aβ: amyloid beta.

[d]GFAP: glial fibrillary acidic protein.

[e]NFL: neurofilament light chain.

[f]Participants with obstructive sleep apnea were included only if they were being treated with continuous positive airway pressure.

[g]MoCA: Montreal Cognitive Assessment. Higher scores indicate better overall cognition.

[h]PSQI: Pittsburgh Sleep Quality Index. Higher scores indicate worse sleep quality.

[i]GAD-7: Generalized Anxiety Disorder-7. Higher scores indicate more symptoms of anxiety.

[j]GDS-15: Geriatric Depression Scale-15 item scale. Higher scores indicate more symptoms of depression.

[k]AES-S: Apathy Evaluation Scale-Self. Higher scores indicate more apathy.

[l]One participant had a missing value for an item, so this score was imputed using simple imputation.

Data Quality and Completeness (Adherence)

Overview

Adherence to the sleep and cognitive tasks is shown in Table 2. Adherence was very high across all study groups and tasks.

Reasons for missing data are summarized in Figure 4 with additional detail provided in Multimedia Appendix 3.

Figure 4. Reasons for missing data are mapped according to the capability, opportunity motivation, behavior (COM-B) model of behavior change.

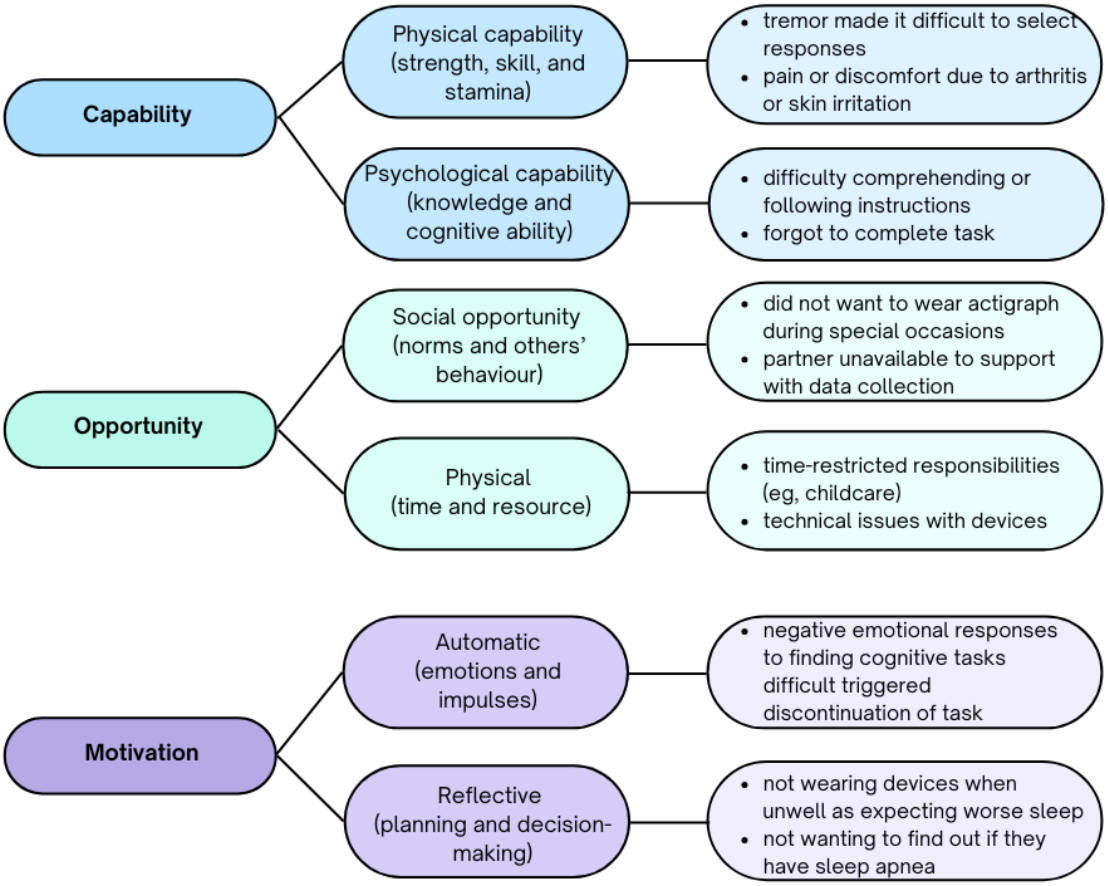


Table 2. A summary of feasibility outcomes relating to data quality and adherence for wearable devices and remote or web-based cognitive assessments.

Study task (device or software); description; and adherence or data quality metric	AD[a] (n=10)	LBD[b] (n=10)	Controls (n=20)	Total (n=40)
Wearable devices				
Sleep EEG[c] (Dreem 2): participants were asked to complete a week of consecutive overnight EEG recordings to record nighttime sleep (7 nights)				
Nights, mean (SD)	5.9 (2)	6.5 (0.7)	6.6 (0.9)	6.4 (1.2)
Full dataset provided (7 nights), n (%)	7 (70)	6 (60)	16 (80)	29 (73)
Record quality[d], mean (SD)	83.5 (12.1)	84.1 (15.0)	85 (18.4)	84.4 (15.8)
Actigraphy (Axivity AX3): participants were asked to complete 24-h wrist actigraphy continuously for 8 weeks (56 nights)				
Nights, mean (SD)	50.3 (13.1)	52.2 (6.8)	54.4 (3.2)	52.8 (7.6)
Full dataset provided (56 nights), n (%)	5 (50)	5 (50)	14 (70)	24 (60)
Pulse oximetry (Nonin WristOx2 3150): participants were asked to complete 2 consecutive nights of overnight pulse oximetry (2 nights)				
Completed ≥1 recording, n (%)	9 (90)	6 (75)	18 (100)	33 (89)
Full dataset provided (2 nights), n (%)[e]	6 (60)	5 (63)	15 (83)	26 (70)
Remote web-based assessments				
Digital sleep diary (MyDignio): participants were prompted to complete daily sleep diaries for 8 weeks (56 nights)				
Diaries completed, mean (SD)	48.2 (13.4)	51.1 (9.8)	53.8 (3.1)	51.7 (8.6)
Full dataset provided (56 nights), n (%)	1 (10)	4 (40)	8 (40)	12 (30)
Web-based cognitive tests (Cognitron): participants were prompted to complete a digit span, choice reaction time (CRT), and self-ordered search (SOS) tasks on twice weekly for 7 weeks and daily for 1 week (21 tasks)				
Digit spans completed, mean (SD)	17.6 (4.3)	17.6 (4.7)	16.8 (4.7)	17.2 (4.5)
CRT[f] completed, mean (SD)	16.9 (4.2)	13.9 (7.5)	16.7 (4.8)	16.1 (5.4)
SOS[g] completed, mean (SD)	17.5 (4.3)	17.5 (4.6)	16.6 (4.7)	17.0 (4.5)
Completed >1 web-based cognitive task, n (%)	10 (100)	8 (80)	20 (100)	38 (95)
Full dataset provided (all 3 cognitive tasks completed across 21 days), n (%)	2 (20)	2 (20)	2 (10)	6 (15)
Verbal memory tasks (videoconferencing): participants were asked to complete 2 evening "learn" tasks and 2 morning recall or recognition tasks scheduled with researchers (4 tasks)				
Memory tasks completed, mean (SD)	3.2 (1.7)	3.4 (1.3)	3.9 (0.4)	3.6 (1.1)
Full dataset provided (4 tasks completed), n (%)	8 (80)	8 (80)	19 (95)	35 (88)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body dementia.

[c]EEG: electroencephalography.

[d]Mean first calculated for each participant across all available recordings.

[e]Three participants who already had a diagnosis of sleep apnea were not invited to complete the pulse oximetry.

[f]CRT: Choice Reaction Time.

[g]SOS: Self-ordered search.

Sleep EEG

In total, 257 recordings were made using Dreem 2. One participant accidentally completed an eighth recording, which was removed from subsequent analyses to reduce bias.

All participants successfully recorded at least 1 night of sleep, with an average of 6.4 (1.2) nights across the cohort or a 91.4% data completeness rate (Table 2). In total, 73% (29/40) of participants provided data for all 7 nights. Average record quality was also high at 84.4% (15.8), indicating nearly optimal record quality on average across the cohort. One participant wore the Dreem 2 but did not successfully initiate the recording during the intensive week but completed 7 successful nights of recording on a second attempt later in the study. Record quality for individual EEG channels is provided in Multimedia Appendix 4.

Actigraphy

In total, 2332 nights of data across the full cohort were collected. A total of 189 of these nights contained no data, either at the

beginning or the end of the actigraphy file and were removed from sleep analyses. Following visual inspection, a further 9 nights were removed due to sustained nonwear. Therefore, 2134 nights were of sufficient quality for sleep analysis.

Across the cohort, participants provided an average of 52.8 (7.6) out of 56 analyzable nights, giving a data completeness rate of 94.3%. In total, 60% (24/40) of participants provided 56 analyzable nights. All participants recorded at least 14 days of actigraphy. Unexpected battery failure outside of participants' control (due to long-term storage without use during the COVID-19 pandemic) affected 2 participants' recordings and was identified during data check at the midpoint study visit. No participants refused to wear or discontinued use of the actigraphy device, although several reported forgetting or choosing not to wear it on certain days or taking brief breaks due to minor skin irritation.

In total, 37 diary entries were manually entered based on visual analysis of the actigraphy data due to missing sleep diaries and poor heuristic algorithm looking at distribution of change in z-angle algorithm detection of the sleep period. Finally, the sleep analysis software was instructed to solely rely on sleep diary information for 166 nights due to a clear misclassification of sleep period.

Overnight Pulse Oximetry

As 3 (8%) out of 40 participants were already being treated for established OSA, 37 (93%) participants were offered overnight pulse oximetry. In total, 3 (8%) participants declined, and 1 (3%) participant was unable to tolerate wearing the device. The remaining 33 (89%) participants completed at least 1 successful overnight recording, with 26 (70%) recording successful oximetry traces on both nights. The data completeness rate was 79.7% for those eligible and asked to complete overnight pulse oximetry. Referrals to a sleep clinic for either sleep apnea or incidental findings (such as abnormal pulse rise index) were indicated for 57% (21/37) of participants. In total, 4 (11%) participants declined referral due to not wanting a formal diagnosis, not wanting to be put on sleep apnea treatment, and the inconvenience of traveling to the clinic. In total, 17 (46%) participants agreed to a referral.

Digital Sleep Diaries

Although few participants (12/40, 30%) completed all 56 sleep diaries, overall adherence was high. Participants completed an average of 51.7 (SD 8.6) sleep diaries, and only 2 (17%) participants completed fewer than 75% of their sleep diaries. The data completeness rate was 92.3%.

Only 1 (3%) participant completed sleep diaries on paper while her partner, who facilitated entries into MyDignio, was away. In total, 3 (8%) participants had technical issues with data entry into the MyDignio app lasting several days due to a software update and provided some of their sleep diaries via email which were then entered by the study team.

At the time of study setup, Dignio did not offer data format validation, and free-text response boxes were used for several questions. Participants often did not input data in the format requested, and inconsistency in reporting prohibited reliable automated recoding; however, most were manually interpretable by the research team. Where a range of values were provided, a mean was calculated and the value was rounded to the nearest whole minute (eg, 5-10 min was recoded to 8 min). One diary entry was manually corrected by the research team to align with other diary entries and the actigraphy file. It was not possible to recode some responses due to ambiguity (eg, responses of "not long," "several minutes," or "unsure"), leaving 43 completed sleep diaries incomplete on at least one sleep variable. Questions most likely to have a missing value were questions on sleep onset latency and length of nocturnal awakenings.

Video-Based Verbal Memory Tasks

In total, 92% (37/40) of participants completed at least 1 verbal memory task with a researcher. Data were missing for 1 (3%) participant as the task was considered inappropriate due to their specific language difficulties, 1 (3%) participant due to technical issues with video calling, and 1 (3%) participant due to distress during the first learning trial. Most participants (35/40, 88%) completed all 4 verbal memory tasks, with 2 (5%) participants completing only 2 of the tasks due to work or childcare commitments. Overall, the data completeness rate was 90%. In total, 3 (8%) participants reported finding the task difficult, 2 (5%) reported feeling distracted during the encoding tasks, 2 (5%) reported finding the encoding task distracting rather than helpful, and 2 (5%) reported finding that usual memory techniques such as the story technique or rehearsing words in an auditory loop were difficult to do in this task.

Web-Based Cognitive Tests

Few participants completed all 21 cognitive tasks: 6 (15%) out of 40 participants completed all 3 cognitive tasks on all 21 occasions. In total, 7 (18%) participants provided complete data for choice reaction time tasks, 8 (20%) for the self-ordered search, and 9 (23%) for the digit span task. However, on average data completeness was good, with 79.8% of allocated web-based cognitive tasks completed. Only 1 (3%) participant with PD attempted the tasks but discontinued due to frustration at not being able to do the tasks, as they required speed and accuracy. Several participants anecdotally reported misunderstanding the choice reaction time task and clicking outside of the response window but these instances were not formally recorded.

Saliva Samples

Overview

The outcomes for saliva and blood samples are summarized in Table 3.

Table 3. The outcomes relating to data quality and adherence for biomarker analysis: blood biomarkers, salivary dim-light melatonin, and salivary cortisol awakening response.

Study task (sample type) and description; adherence or data quality metric	AD[a] (n=10)	LBD[b] (n=10)	Controls (n=20)	Total (n=40)
Neurodegenerative and AD biomarker samples (blood): participants were asked to have blood drawn for biomarker testing				
Participants who provided blood for biomarker testing, n (%)	9 (90)	10 (100)	18 (90)	37 (92)
Samples analyzed for all biomarkers, n (%)	8 (80)	8 (80)	14 (70)	30 (75)
Samples analyzed for ≥1 biomarker, n (%)	9 (90)	10 (100)	18 (90)	37 (92)
Melatonin samples (passive drool): participants were asked to complete 7 hourly samples across 1 evening, starting from 5 h before bed, and record the sample time				
Analyzable samples provided per participant, mean (SD)	6.3 (2.2)	6.9 (0.3)	7.0 (0)	6.8 (1.1)
Participants who completed all 7 samples, n (%)	9 (90)	9 (90)	20 (100)	38 (95)
Average discrepancy between scheduled time and reported time of completion (min)[c], mean (SD)	5.6 (53.4)	−13.9 (32.6)	−9.2 (37.1)	−7.0 (41)
Cortisol samples (oral swab): participants were asked to complete 3 samples across 1 morning: upon waking up, after 30 min, and after 60 min, and record the sample time				
Analyzable samples provided per participant, mean (SD)	2.7 (0.9)	3.0 (0)	3.0 (0)	2.9 (0.5)
Participants who completed all 3 samples, n (%)	9 (90)	10 (100)	20 (100)	39(98)
Average discrepancy between scheduled time and time completed (min)[c], mean (SD)	22.5 (27.8)	24.3 (17)	24.6 (23.8)	24.0 (23.2)

[a]AD: Alzheimer disease.

[b]LBD: Lewy body dementia.

[c]Positive numbers here refer to a sample being taken later than scheduled, and negative numbers refer to a sample being taken earlier than scheduled.

Dim-Light Melatonin Assay: Passive Drool Samples

Most saliva samples (272/280, 97.1%) were collected as requested. Of 40 participants, 1 (3%) participant provided no saliva samples, and 1 (3%) participant provided 9 out of 10 samples. Data completeness rate for analyzable samples was 92.5% (259/280) across the cohort. Participant-reported sample times were available for 86.7% (236/272) of the total sample. In total, 50.8% (120/236) of the total sample were reported to have been completed within 15 minutes of the scheduled time, as calculated by average bedtimes reported in the sleep diary. On average, participants completed their samples around 7 minutes early (Table 3). Previous studies have identified highly variable salivary melatonin secretion, with peaks from 2 pg/mL to 84 pg/mL [57]. Across the cohort, values ranged from beneath detectable levels (<1.37 pg/mL) to 92.4 pg/mL.

Cortisol Awakening Response: Oral Saliva Swabs

Participants were instructed to provide saliva samples upon waking up, after 30 minutes, and after 60 minutes on one morning during the intensive week. Data completeness was 97.5%, with missing data from only 1 participant who did not provide any saliva samples. Data completeness rate for analyzable samples was 91.7% across the cohort. We compared reported saliva sample timings to their scheduled timings, based on final awakening time estimated by EEG. Participant-reported sample times and EEG awakening time were available for 72.6% (85/117) of cortisol samples. From the 85 samples with timings available, suggested wake times were later than the first sample recording time for 6 samples (7%) and were removed from analysis due to suspected error in recorded saliva timing or date. Participants recorded their first cortisol saliva swab an average of 23 minutes after awakening and overall cortisol samples were generally 24 minutes later than scheduled across the 3 time points (Table 3). Previous studies have identified variable salivary cortisol awakening values from 3 to 19 µg/L in healthy adults [58]. Across the cohort, values ranged from 0.9 to 13.4 µg/L.

Associations Between Participant Characteristics and Adherence

There was some evidence of a weak correlation between MoCA score at baseline and EEG record quality, where those with greater cognitive impairment had lower average EEG record quality (r=0.31; *P*=.05). Spearman rank correlations revealed no evidence of correlations between EEG record quality and age (r=−0.10; *P*=.54); apathy (r=0.16; *P*=.31); or PSQI score (r=0.19; *P*=.24). There was also no evidence of correlation between the number of sleep diaries completed and age (r=−0.15; *P*=.35); MoCA score (r=0.11; *P*=.48); apathy (r=0.07; *P*=.69); or PSQI score (r=0.12; *P*=.47).

Resource Use

Support From Partners and Relatives

In total, 36% (4/11) of individuals with AD, 85% (11/13) of individuals with PD or LBD, and 10% (2/20) of controls attended their consent visit with a partner or relative. In total, 80% (32/40) of participants reported that there would be someone external to the research team (eg, partner or relative) who could support them with study tasks if needed at baseline. However, only 30% (12/40) of participants reported receiving support with completing study tasks (n=3, 30% AD; n=8, 80% LBD; and n=1, 5% control). Support from outside of the study

team was predominantly from partners and included reminders to complete tasks, setting up devices, and troubleshooting technical problems. Of 40 participants, 3 (8%) participants who required technical support from a partner or relative, or used their devices to complete study tasks, reported missing study tasks due to their partner or relative being unavailable during the study period. The participant who withdrew following baseline reported having someone who could support them at home.

Support and Contact With the Research Team

Participants received in person and remote training and support from a researcher to complete study tasks at baseline (including downloading and setting up apps and turning on recording devices) and were offered a refresher training session before the intensive week. Participants were also provided with a written instruction manual, scheduled in-app and email reminders, and ad hoc support as requested by participants or where the team noticed ≥3 consecutive days of missing sleep diary data. All participants had at least 1 in-person visit from a researcher (eg, to download midpoint actigraphy data or retrieve study equipment) during the study period. Participants were advised to contact the research team if they had questions during the study by their preferred method (email, phone, video call, or instant messaging on the MyDignio app). Support from the research team during the study was most often provided by email and involved responding to participant queries on initial setup (eg, downloading the MyDignio app); reminders (eg, tasks to complete or passwords); and technical issues (eg, links not arriving for Cognitron). Some participants also requested telephone-based, video-based, or home-based support (eg, to refresh training on the intensive week study tasks). Most participants did not receive regular reminders from the study team to complete tasks.

Device Use

One participant was provided with a study tablet upon request, as they wanted to keep study activities and apps separate to their personal devices. All other participants used their own smart devices for study activities (ie, smartphones, tablets, and personal computers). Most participants used >1 device to complete web-based study tasks due to personal preference or convenience. No devices provided by the study team were lost or damaged during the study, although some participants cut the actigraphy watch strap for comfort.

Discussion

Principal Findings

In this study, we show that it is feasible to remotely measure sleep and cognition longitudinally in community-dwelling older adults with MCI and dementia due to AD and LBD and healthy older adults. Eligible participants were interested in, enrolled in, and remained in the study, despite being asked to complete a high volume of remote and novel study tasks across an 8-week period while continuing their usual routines. Only 1 participant withdrew from the study, and this was before remote data collection started. Most participants were receptive to multimodal home-based research using technology and alternatives to in-laboratory sleep assessment. Across the cohort, data completeness rate was high and ranged from 79.8% to 97.1%. Our findings support the use of remote, technology-supported research methods to study natural sleep in future trials and indicate some areas where further improvements and refinements would be helpful. With just under two-thirds of our sample (25/40, 63%) having sleep apnea, and only 3 (8%) being aware of this before joining the study, our results also highlight the importance of sleep apnea screening in older adults, as a risk factor and potentially reversible contributor to cognitive impairment and poor sleep quality [59].

Many older adults with MCI and dementia routinely leverage technology for cognitive stimulation, performing activities of daily living (such as shopping and banking), entertainment, and socializing, and use assistive-technology solutions, such as reminders and navigation to support independence [60]. However, cognitive impairment might impact ability to understand or remember to complete remote or technology-based study tasks, resulting in missing data, and may trigger anxiety if patients feel unable to perform tasks correctly [61]. Although a few previous studies had examined the use of sleep wearables in adults with AD, this was typically done over a brief period, examined feasibility of using only 1 digital health technology, or required input from a carer or study partner [31-34,62]. Our study demonstrated that older adults with MCI and dementia can successfully complete novel multimodal sleep and cognitive assessments, including longitudinal concurrent use of wrist actigraphy, web-based cognitive tests, digital sleep diaries, and wireless EEG headbands. However, more effort will be required to recruit samples which are representative of the older adult population and those with MCI or dementia, as our study predominantly recruited White male individuals who were familiar with smart technology, albeit not the devices used in RESTED.

Data completeness rates, reflecting both the ability to complete study tasks and produce data of analyzable quality, were high across all study tasks and participants. Remote data collection comes with a risk of missing data or nonadherence; however, we were able to mitigate this risk by using devices which upload data to servers at regular intervals (eg, Dreem and Dignio) and, less conveniently, by downloading data manually from devices to check compliance and data quality and offer support where needed (eg, actigraphy). Passive monitoring devices and devices that upload to servers automatically can simultaneously reduce the burden on participants and researchers while minimizing data loss by allowing researchers to monitor and respond quickly to any user or technical issues.

Reasons for missing data were usually known and typically related to infrequent but intentional decisions to remove a device (eg, due to a social event or illness) or, in most cases, technical issues with devices or software beyond the participants' control, which would not introduce bias. We observed a weak correlation between EEG record quality and baseline cognitive impairment, which could be explained by cognitive impairment impacting ability to correctly wear the EEG headband for maximal impedance or reflect that patients were moving more during the night. Even subtle movements can shift the headband and reduce

signal quality and getting optimal signal quality across entire or several recordings is a recognized challenge for dry EEG [52]. More comprehensive cognitive assessment than the MoCA, particularly a deeper assessment of executive dysfunction and long-term memory, might reveal a stronger relationship between cognitive impairment and adherence. However, recent studies using remote monitoring technologies found that, although patients with more advanced neurodegenerative disease or cognitive impairment reported more problems and were less compliant, the study remained feasible [63,64]. Larger studies may also want to consider regression analyses to identify key predictors of adherence to identify where support might be most indicated by the research team or caregiver support. However, importantly, overall record quality remained high across the cohort. Having MCI or mild dementia did not preclude study participation, adherence to remote supervised and unsupervised study tasks, or obtaining good-quality data.

Use of remote sleep and cognitive monitoring can help to partly or fully decentralize research, as participants can use their own devices or be provided with technology via post where needed [65], which is more scalable for clinical trials [66]. Frequent or lengthy clinic visits, particularly those that require overnight stays for sleep analysis, can be burdensome for participants, and may reduce inclusivity by requiring participants to live near a study site or travel long distances. More frequent but briefer remote assessments may also help to mitigate against the risk of large amounts of missing data resulting from a missed clinic appointment, can act as “digital biomarkers” that may detect changes more sensitively than an annual follow-up [67], and are more convenient for participants [68]. Although not the focus of our work, research could also explore the use of digital tools to assess fluctuations in cognition over the day in LBD (where fluctuation in cognition is a key diagnostic criteria) and AD (where a sundowning effect of heightened distress is often observed during the evening hours), both of which are difficult to assess in clinical settings [69,70].

While high adherence has been previously demonstrated in feasibility studies of remote monitoring technologies in participants with cognitive impairment or neurodegenerative conditions, most studies have required intensive study partner support throughout the study period as an eligibility criterion [63,71]. It has been argued that study partners are essential for participant safety and well-being in dementia research, even at the preclinical stage where cognitively normal individuals are enrolled [72]. For some participants, partners or relatives provided essential support for study participation. However, many participants provided good-quality data with minimal to no input from others. Crucially, several participants would not have been able to participate if there had been a study partner requirement. Strict eligibility criteria for trials are a recognized barrier to research participation [73]. Requiring a study partner for all studies may unnecessarily increase the burden for relatives already providing care or support, reducing the participant pool, and may disproportionately affect different groups in society who are underserved by research [74,75]. Our findings encourage researchers to consider inclusive and flexible research designs, including options to formally recruit and collect data from study partners without excluding participants who do not have someone available to support them in research studies.

There are an increasing number of options when considering how to measure sleep and cognition from home, including both research-grade and consumer sleep trackers [76,77]. We opted to use existing technologies that were noninvasive, affordable, required minimal training or supervision, were commercially available research-grade or consumer-grade devices that met necessary data privacy regulations, and provided data in a format that could be analyzed using open-source software. Existing technologies can be implemented immediately, reducing time and cost to setup research studies, have a more mature user interface, and are less error-prone than a newly developed solution, and crucially, may be easier to compare among studies for meta-analysis. Several study tasks required ongoing technical support or services from the manufacturers or developers. As a customer rather than a collaborator, we could not always identify or respond to technical problems, and with some providers, we were not informed in advance of several significant changes, which impacted the study or its participants resulting in data loss. Collaborating with industry, ideally at an early stage of the research process, might ensure longevity and continued support throughout the life course of the study and increased flexibility to adapt technology to better fit research (eg, data validation for digital patient-reported outcomes).

Limitations

Women, adults aged >80 years, and minoritized ethnic groups were underrepresented in our study, which is often observed in dementia research [78]. Although some barriers (eg, mistrust and accessibility of research) and motivators (eg, altruism) have been identified, further work is needed to identify how to improve research access, inclusion, and participation in dementia research [79]. There may also have been a self-selection bias, whereby individuals who were likely to be more competent and enthusiastic about using technology volunteered to participate in the study. Most of our participants were regular smartphone users, and several prospective participants declined to participate due to concerns regarding their confidence or interest in using the technologies. Although the proportion of older adults in the United Kingdom using the internet is increasing, a significant proportion of older adults do not regularly use the internet or lack fundamental skills such as being able to turn on a device and enter login details [28]. Older adults with MCI and dementia who regularly use devices report that smartphones and tablets can be useful to support activities of daily living, as well as for communication, entertainment, and recreation, but also list concerns, including cybersecurity and vulnerability to fraud [80]. Education around sleep and brain health, and basic digital skills training on how devices can support individuals living with cognitive impairment, may improve perceptions of capability and motivation to participate in similar studies [56,80]. Exploring the feasibility of more passive monitoring technologies, such as mattress sensors or smartphone-based passive sensing, may help to increase participation and confidence from those with less experience or interest in technology. Although these techniques may invite additional concerns around data privacy, passive monitoring technologies are already being used to support older adults aging in place

[81,82]. Increasing the pool of prospective participants through recruiting via multiple sites and meaningful engagement with community groups, and considering digital inclusivity is likely to enhance recruitment in future studies. Finally, we did not set a priori feasibility cutoffs for study tasks. Despite these limitations, the high completion and retention rate from the RESTED study suggests that remote sleep and cognitive monitoring could offer detailed sleep profiling suitable for tracking change in sleep over disease progression or for monitoring change in sleep clinical trials in older adults with or at risk of dementia.

Conclusions

Practical, detailed, and scalable assessment of sleep is essential for understanding how sleep disturbances affect neurodegeneration over time and for developing and evaluating effective interventions. Our results suggest that older adults with MCI and dementia, as well as healthy older adults, can and do engage in multimodal remote sleep and cognitive research, including using wireless EEG, actigraphy, and mobile apps or web applications. Remote monitoring technologies and research designs offer the opportunity to study natural sleep and its relationship with cognition and dementia over extended periods, are scalable, and should be considered when designing future clinical trials in sleep and dementia in these populations.

Acknowledgments

The authors sincerely thank the Remote Evaluation of Sleep to Enhance Understanding of Early Dementia (RESTED) study participants and their relatives and partners who supported them to take part. They would also like to thank the ReMemBr group lived experience experts for their contributions to improving the study design and supporting them to make the study as accessible as possible. They would also like to thank the North Bristol NHS Trust Respiratory Physiology department and the Bristol Bioresource Laboratories, Anglia Ruskin University Biomarker Lab, and UK DRI Fluid Biomarker Laboratory for providing their services and support for our saliva and blood biomarker analyses.

Finally, the authors would like to thank the teams at Dreem Research, Cognitron, Join Dementia Research, and Ewa G Truchanowicz and Phil Reay, and their teams at Dignio, United Kingdom.

The RESTED study was funded by charitable organizations (BRACE, Alzheimer's Research UK, the Bristol and Weston Hospitals Charity [previously Above and Beyond], and the David Telling Charitable Trust), the NIHR Bristol Biomedical Research Centre, and philanthropic donations from S Scobie and A Graham. JB also received funding from Alzheimer's Research United Kingdom (ARUK) supported by the Margaret Jost Fellowship and the Don Thoburn Memorial Scholarship. The funders had no role in the trial design, patient recruitment, data collection, analysis, interpretation, writing of the manuscript, or the decision to submit it for publication.

Data Availability

Data are available from the author upon reasonable request.

Authors' Contributions

VGG contributed to conceptualization, methodology, software, project administration, resources, investigation, data curation, formal analysis, visualization, writing—original draft preparation, and writing—review and editing. JB contributed to conceptualization, methodology, software, formal analysis, investigation, writing—review and editing, and funding acquisition. HM contributed to conceptualization, methodology, investigation, writing—review and editing, and funding acquisition. HL contributed to investigation, project administration, and formal analysis. AK contributed to conceptualization, methodology, and formal analysis. NT contributed to methodology, funding acquisition, and writing—review and editing. RG contributed to conceptualization, funding acquisition. BB contributed to methodology. EC contributed to conceptualization, methodology, resources, writing—review and editing, supervision, and funding acquisition.

Conflicts of Interest

EC has received funding from Biogen, Eisai, and Lilly for consultancy and providing educational resources. AH has received funding from Quanterix Corp for consultancy. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) Statement—checklist of items that should be included in reports of observational studies.
[DOCX File , 19 KB - aging_v8i1e72824_app1.docx]

Multimedia Appendix 2
Further information is provided on the devices used to measure sleep and circadian rhythms.

[DOCX File , 13 KB - aging_v8i1e72824_app2.docx]

Multimedia Appendix 3
Dreem electroencephalography individual channel metrics.
[DOCX File , 14 KB - aging_v8i1e72824_app3.docx]

Multimedia Appendix 4
Additional information on reasons for missing data and the affected study tasks is provided.
[DOCX File , 15 KB - aging_v8i1e72824_app4.docx]

References

1. Koren T, Fisher E, Webster L, Livingston G, Rapaport P. Prevalence of sleep disturbances in people with dementia living in the community: a systematic review and meta-analysis. Ageing Res Rev 2023 Jan;83:101782 [FREE Full text] [doi: 10.1016/j.arr.2022.101782] [Medline: 36356799]
2. Postuma RB, Iranzo A, Hu M, Högl B, Boeve BF, Manni R, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain 2019 Mar 01;142(3):744-759 [FREE Full text] [doi: 10.1093/brain/awz030] [Medline: 30789229]
3. Harenbrock J, Holling H, Reid G, Koychev I. A meta-analysis of the relationship between sleep and β-Amyloid biomarkers in Alzheimer's disease. Biomark Neuropsychiatry 2023 Dec;9:100068. [doi: 10.1016/j.bionps.2023.100068]
4. Zamore Z, Veasey SC. Neural consequences of chronic sleep disruption. Trends Neurosci 2022 Sep;45(9):678-691 [FREE Full text] [doi: 10.1016/j.tins.2022.05.007] [Medline: 35691776]
5. Zielinski MR, Gibbons AJ. Neuroinflammation, sleep, and circadian rhythms. Front Cell Infect Microbiol 2022 Mar 22;12:853096 [FREE Full text] [doi: 10.3389/fcimb.2022.853096] [Medline: 35392608]
6. Eshera YM, Gavrilova L, Hughes JW. Sleep is essential for cardiovascular health: an analytic review of the relationship between sleep and cardiovascular mortality. Am J Lifestyle Med 2024 Oct 31;18(3):340-350. [doi: 10.1177/15598276231211846] [Medline: 38737888]
7. Seda G, Matwiyoff G, Parrish JS. Effects of obstructive sleep apnea and CPAP on cognitive function. Curr Neurol Neurosci Rep 2021 May 06;21(7):32. [doi: 10.1007/s11910-021-01123-0] [Medline: 33956247]
8. Costa YS, Lim AS, Thorpe KE, Colelli DR, Mitchell S, Masellis M, et al. Investigating changes in cognition associated with the use of CPAP in cognitive impairment and dementia: a retrospective study. Sleep Med 2023 Jan;101:437-444. [doi: 10.1016/j.sleep.2022.11.037] [Medline: 36516600]
9. Liu WT, Huang HT, Hung HY, Lin SY, Hsu WH, Lee FY, et al. Continuous positive airway pressure reduces plasma neurochemical levels in patients with OSA: a pilot study. Life (Basel) 2023 Feb 22;13(3):613 [FREE Full text] [doi: 10.3390/life13030613] [Medline: 36983769]
10. Sabia S, Fayosse A, Dumurgier J, van Hees VT, Paquet C, Sommerlad A, et al. Association of sleep duration in middle and old age with incidence of dementia. Nat Commun 2021 Apr 20;12(1):2289 [FREE Full text] [doi: 10.1038/s41467-021-22354-2] [Medline: 33879784]
11. Mayer G, Frohnhofen H, Jokisch M, Hermann DM, Gronewold J. Associations of sleep disorders with all-cause MCI/dementia and different types of dementia - clinical evidence, potential pathomechanisms and treatment options: a narrative review. Front Neurosci 2024 Mar 22;18:1372326 [FREE Full text] [doi: 10.3389/fnins.2024.1372326] [Medline: 38586191]
12. Blackman J, Swirski M, Clynes J, Harding S, Leng Y, Coulthard E. Pharmacological and non-pharmacological interventions to enhance sleep in mild cognitive impairment and mild Alzheimer's disease: a systematic review. J Sleep Res 2021 Aug;30(4):e13229 [FREE Full text] [doi: 10.1111/jsr.13229] [Medline: 33289311]
13. Blackman J, Morrison HD, Lloyd K, Gimson A, Banerjee LV, Green S, et al. The past, present, and future of sleep measurement in mild cognitive impairment and early dementia-towards a core outcome set: a scoping review. Sleep 2022 Jul 11;45(7):3-11 [FREE Full text] [doi: 10.1093/sleep/zsac077] [Medline: 35373837]
14. Kainec KA, Caccavaro J, Barnes M, Hoff C, Berlin A, Spencer RM. Evaluating accuracy in five commercial sleep-tracking devices compared to research-grade actigraphy and polysomnography. Sensors (Basel) 2024 Jan 19;24(2):635 [FREE Full text] [doi: 10.3390/s24020635] [Medline: 38276327]
15. Tadokoro K, Ohta Y, Hishikawa N, Nomura E, Wakutani Y, Takao Y, et al. Discrepancy of subjective and objective sleep problems in Alzheimer's disease and mild cognitive impairment detected by a home-based sleep analysis. J Clin Neurosci 2020 Apr;74:76-80. [doi: 10.1016/j.jocn.2020.01.085] [Medline: 32007378]
16. Casagrande M, Forte G, Favieri F, Corbo I. Sleep quality and aging: a systematic review on healthy older people, mild cognitive impairment and Alzheimer's disease. Int J Environ Res Public Health 2022 Jul 11;19(14):8457 [FREE Full text] [doi: 10.3390/ijerph19148457] [Medline: 35886309]
17. Williams JM, Kay DB, Rowe M, McCrae CS. Sleep discrepancy, sleep complaint, and poor sleep among older adults. J Gerontol B Psychol Sci Soc Sci 2013 Sep 26;68(5):712-720 [FREE Full text] [doi: 10.1093/geronb/gbt030] [Medline: 23804432]

18. Markun LC, Sampat A. Clinician-focused overview and developments in polysomnography. Curr Sleep Med Rep 2020 Nov 23;6(4):309-321 [FREE Full text] [doi: 10.1007/s40675-020-00197-5] [Medline: 33251088]
19. Goparaju B, de Palma G, Bianchi MT. Naturalistic sleep tracking in a longitudinal cohort: how long is long enough? medRxiv Preprint posted online October 21, 2024 [FREE Full text] [doi: 10.1101/2024.10.19.24315818]
20. Blackman J, Morrison H, Lloyd K, Gimson A, Banerjee L, Green S, et al. Sleep measurement heterogeneity in mild cognitive impairment and early dementia - towards a core outcome set: a scoping review. Sleep 2022 Dec;100:S148-S149. [doi: 10.1016/j.sleep.2022.05.403]
21. Green SF, Frame T, Banerjee LV, Gimson A, Blackman J, Morrison H, et al. A systematic review of the validity of non-invasive sleep-measuring devices in mid-to-late life adults: future utility for Alzheimer's disease research. Sleep Med Rev 2022 Oct;65:101665 [FREE Full text] [doi: 10.1016/j.smrv.2022.101665] [Medline: 36208588]
22. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019 Feb 21;2(1):9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
23. Arnal PJ, Thorey V, Debellemaniere E, Ballard ME, Bou Hernandez A, Guillot A, et al. The Dreem Headband compared to polysomnography for electroencephalographic signal acquisition and sleep staging. Sleep 2020 Nov 12;43(11):zsaa097 [FREE Full text] [doi: 10.1093/sleep/zsaa097] [Medline: 32433768]
24. Miller DJ, Sargent C, Roach GD. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. Sensors (Basel) 2022 Aug 22;22(16):6317 [FREE Full text] [doi: 10.3390/s22166317] [Medline: 36016077]
25. Cay G, Ravichandran V, Sadhu S, Zisk AH, Salisbury AL, Solanki D, et al. Recent advancement in sleep technologies: a literature review on clinical standards, sensors, apps, and AI methods. IEEE Access 2022;10:104737-104756. [doi: 10.1109/access.2022.3210518]
26. Ding Z, Lee T, Chan AS. Digital cognitive biomarker for mild cognitive impairments and dementia: a systematic review. J Clin Med 2022 Jul 19;11(14):4191 [FREE Full text] [doi: 10.3390/jcm11144191] [Medline: 35887956]
27. Masanneck L, Gieseler P, Gordon WJ, Meuth SG, Stern AD. Evidence from ClinicalTrials.gov on the growth of digital health technologies in neurology trials. NPJ Digit Med 2023 Feb 10;6(1):23 [FREE Full text] [doi: 10.1038/s41746-023-00767-1] [Medline: 36765123]
28. Briefing: facts and figures about digital inclusion and older people. Age UK. URL: https://www.ageuk.org.uk/siteassets/documents/reports-and-publications/reports-and-briefings/active-communities/internet-use-statistics-june-2024.pdf [accessed 2024-05-29]
29. Holthe T, Halvorsrud L, Karterud D, Hoel K, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886 [FREE Full text] [doi: 10.2147/CIA.S154717] [Medline: 29765211]
30. Chien SY, Zaslavsky O, Berridge C. Technology usability for people living with dementia: concept analysis. JMIR Aging 2024 Jul 03;7:e51987 [FREE Full text] [doi: 10.2196/51987] [Medline: 38959053]
31. Kent BA, Casciola AA, Carlucci SK, Chen M, Stager S, Mirian MS, et al. Home EEG sleep assessment shows reduced slow-wave sleep in mild-moderate Alzheimer's disease. Alzheimers Dement (N Y) 2022 Aug 16;8(1):e12347 [FREE Full text] [doi: 10.1002/trc2.12347] [Medline: 35992215]
32. Pavlickova H, Russell AE, Lightman S, McCabe R. Feasibility of salivary cortisol collection in patients and companions attending dementia diagnostic meetings in memory clinics. BMC Res Notes 2021 Jan 21;14(1):30 [FREE Full text] [doi: 10.1186/s13104-021-05446-6] [Medline: 33478564]
33. Jones C, Moyle W. A feasibility study of Dreampad™ on sleep, wandering and agitated behaviors in people living with dementia. Geriatr Nurs 2020 Nov;41(6):782-789. [doi: 10.1016/j.gerinurse.2020.04.014] [Medline: 32522427]
34. Guu TW, Brem A, Albertyn CP, Kandangwa P, Aarsland D, Ffytche D. Wrist-worn actigraphy in agitated late-stage dementia patients: a feasibility study on digital inclusion. Alzheimers Dement 2024 May 18;20(5):3211-3218 [FREE Full text] [doi: 10.1002/alz.13772] [Medline: 38497216]
35. Gabb VG, Blackman J, Morrison HD, Biswas B, Li H, Turner N, et al. Remote evaluation of sleep and circadian rhythms in older adults with mild cognitive impairment and dementia: protocol for a feasibility and acceptability mixed methods study. JMIR Res Protoc 2024 Mar 22;13:e52652 [FREE Full text] [doi: 10.2196/52652] [Medline: 38517469]
36. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct;370(9596):1453-1457. [doi: 10.1016/s0140-6736(07)61602-x]
37. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
38. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]

39. Yamada M, Komatsu J, Nakamura K, Sakai K, Samuraki-Yokohama M, Nakajima K, et al. Diagnostic criteria for dementia with Lewy bodies: updates and future directions. J Mov Disord 2020 Jan;13(1):1-10 [FREE Full text] [doi: 10.14802/jmd.19052] [Medline: 31694357]
40. Litvan I, Goldman JG, Tröster AI, Schmand BA, Weintraub D, Petersen RC, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines. Mov Disord 2012 Mar 24;27(3):349-356 [FREE Full text] [doi: 10.1002/mds.24893] [Medline: 22275317]
41. Blackman J, Morrison HD, Gabb V, Biswas B, Li H, Turner N, et al. Remote evaluation of sleep to enhance understanding of early dementia due to Alzheimer's Disease (RESTED-AD): an observational cohort study protocol. BMC Geriatr 2023 Sep 23;23(1):590 [FREE Full text] [doi: 10.1186/s12877-023-04288-0] [Medline: 37742001]
42. Blackman J, Gabb VG, Carrigan N, Wearn A, Meky S, Selwood J, et al. Sleep quality during and after severe acute respiratory syndrome coronavirus 2 (COVID-19) lockdowns in the UK: results from the SleepQuest study. J Sleep Res 2024 Dec 23;33(6):e14205. [doi: 10.1111/jsr.14205] [Medline: 38650540]
43. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
44. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
45. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [doi: 10.1093/sleep/14.6.540] [Medline: 1798888]
46. Chung F, Abdullah HR, Liao P. STOP-bang questionnaire: a practical approach to screen for obstructive sleep apnea. Chest 2016 Mar;149(3):631-638 [FREE Full text] [doi: 10.1378/chest.15-0903] [Medline: 26378880]
47. Yesavage JA, Sheikh JI. Geriatric Depression Scale (GDS): recent findings and development of a shorter version. In: Brink TL, editor. Clinical Gerontology: A Guide to Assessment and Intervention. Oxfordshire, UK: Routledge; Oct 25, 2008:165-173.
48. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
49. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the apathy evaluation scale. Psychiatry Res 1991 Aug;38(2):143-162. [doi: 10.1016/0165-1781(91)90040-v] [Medline: 1754629]
50. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
51. Cognitron. URL: https://www.cognitron.co.uk/ [accessed 2025-05-29]
52. Ravindran KKG, Della Monica C, Atzori G, Nilforooshan R, Hassanin H, Revell V, et al. Evaluation of Dreem headband for sleep staging and EEG spectral analysis in people living with Alzheimer's and older adults. Sleep 2025 May 04:9240. [doi: 10.1093/sleep/zsaf122] [Medline: 40319389]
53. González DA, Wang D, Pollet E, Velarde A, Horn S, Coss P, et al. Performance of the Dreem 2 EEG headband, relative to polysomnography, for assessing sleep in Parkinson's disease. Sleep Health 2024 Feb;10(1):24-30. [doi: 10.1016/j.sleh.2023.11.012] [Medline: 38151377]
54. Migueles JH, Rowlands AV, Huber F, Sabia S, van Hees VT. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. J Meas Phys Behav 2022 Jul 01;31(5):629 [FREE Full text] [doi: 10.1123/jmpb.2018-0063]
55. van Hees VT, Sabia S, Jones SE, Wood AR, Anderson KN, Kivimäki M, et al. Estimating sleep parameters using an accelerometer without sleep diary. Sci Rep 2018 Aug 28;8(1):12975 [FREE Full text] [doi: 10.1038/s41598-018-31266-z] [Medline: 30154500]
56. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
57. Burgess HJ, Fogg LF. Individual differences in the amount and timing of salivary melatonin secretion. PLoS One 2008 Aug 26;3(8):e3055 [FREE Full text] [doi: 10.1371/journal.pone.0003055] [Medline: 18725972]
58. Trilck M, Flitsch J, Lüdecke DK, Jung R, Petersenn S. Salivary cortisol measurement--a reliable method for the diagnosis of Cushing's syndrome. Exp Clin Endocrinol Diabetes 2005 Apr 13;113(4):225-230. [doi: 10.1055/s-2005-837667] [Medline: 15891959]
59. Naismith S, Winter V, Gotsopoulos H, Hickie I, Cistulli P. Neurobehavioral functioning in obstructive sleep apnea: differential effects of sleep quality, hypoxemia and subjective sleepiness. J Clin Exp Neuropsychol 2004 Feb 09;26(1):43-54. [doi: 10.1076/jcen.26.1.43.23929] [Medline: 14972693]
60. Holthe T, Halvorsrud L, Lund A. Digital assistive technology to support everyday living in community-dwelling older adults with mild cognitive impairment and dementia. Clin Interv Aging 2022;17:519-544 [FREE Full text] [doi: 10.2147/CIA.S357860] [Medline: 35464157]

61. Ahuja M, Siddhpuria S, Reppas-Rindlisbacher C, Wong E, Gormley J, Lee J, et al. Sleep monitoring challenges in patients with neurocognitive disorders: a cross-sectional analysis of missing data from activity trackers. Health Sci Rep 2022 May 26;5(3):e608 [FREE Full text] [doi: 10.1002/hsr2.608] [Medline: 35509396]
62. Van Den Berg JF, Van Rooij FJ, Vos H, Tulen JH, Hofman A, Miedema HM, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. J Sleep Res 2008 Sep 21;17(3):295-302. [doi: 10.1111/j.1365-2869.2008.00638.x] [Medline: 18321246]
63. Muurling M, de Boer C, Hinds C, Atreya A, Doherty A, Alepopoulos V, RADAR-AD consortium. Feasibility and usability of remote monitoring in Alzheimer's disease. Digit Health 2024 Apr 09;10:20552076241238133 [FREE Full text] [doi: 10.1177/20552076241238133] [Medline: 38601188]
64. Maas BR, Speelberg DH, de Vries G, Valenti G, Ejupi A, Bloem BR, et al. Patient experience and feasibility of a remote monitoring system in Parkinson's disease. Mov Disord Clin Pract 2024 Oct 26;11(10):1223-1231. [doi: 10.1002/mdc3.14169] [Medline: 39056543]
65. Aiyegbusi OL, Davies EH, Myles P, Williams T, Frost C, Haroon S, et al. Digitally enabled decentralised research: opportunities to improve the efficiency of clinical trials and observational studies. BMJ Evid Based Med 2023 Oct 21;28(5):328-331 [FREE Full text] [doi: 10.1136/bmjebm-2023-112253] [Medline: 36810190]
66. Inan OT, Tenaerts P, Prindiville SA, Reynolds HR, Dizon DS, Cooper-Arnold K, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3(1):101. [doi: 10.1038/S41746-020-0302-Y]
67. Dodge HH, Zhu J, Mattek NC, Austin D, Kornfeld J, Kaye JA. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. PLoS One 2015 Sep 17;10(9):e0138095 [FREE Full text] [doi: 10.1371/journal.pone.0138095] [Medline: 26379170]
68. Black BS, Taylor HA, Rabins PV, Karlawish J. Study partners perform essential tasks in dementia research and can experience burdens and benefits in this role. Dementia (London) 2016 May 13;17(4):494-514 [FREE Full text] [doi: 10.1177/1471301216648796] [Medline: 27179001]
69. O'Dowd S, Schumacher J, Burn DJ, Bonanni L, Onofrj M, Thomas A, et al. Fluctuating cognition in the Lewy body dementias. Brain 2019 Nov 01;142(11):3338-3350. [doi: 10.1093/brain/awz235] [Medline: 31411317]
70. Canevelli M, Valletta M, Trebbastoni A, Sarli G, D'Antonio F, Tariciotti L, et al. Sundowning in dementia: clinical relevance, pathophysiological determinants, and therapeutic approaches. Front Med (Lausanne) 2016 Dec 27;3:73 [FREE Full text] [doi: 10.3389/fmed.2016.00073] [Medline: 28083535]
71. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May 27;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
72. Grill JD, Karlawish J. Study partners should be required in preclinical Alzheimer's disease trials. Alzheimers Res Ther 2017 Dec 06;9(1):93 [FREE Full text] [doi: 10.1186/s13195-017-0327-x] [Medline: 29212555]
73. Mitchell AK, Ehrenkranz R, Franzen S, Han SH, Shakur M, McGowan M, et al. Analysis of eligibility criteria in Alzheimer's and related dementias clinical trials. Sci Rep 2024 Jul 01;14(1):15036 [FREE Full text] [doi: 10.1038/s41598-024-65767-x] [Medline: 38951633]
74. Aranda MP, Marquez DX, Gallagher-Thompson D, Pérez A, Rojas JC, Hill CV, et al. A call to address structural barriers to Hispanic/Latino representation in clinical trials on Alzheimer's disease and related dementias: a micro-meso-macro perspective. Alzheimers Dement (N Y) 2023;9(2):e12389 [FREE Full text] [doi: 10.1002/trc2.12389] [Medline: 37287471]
75. Wise-Brown A, Brangman SA, Henderson JN, Willis-Parker M, Monroe S, Mintzer JE, et al. Promoting diversity in clinical trials: insights from planning the ALUMNI AD study in historically underrepresented US populations with early symptomatic Alzheimer's disease. EClinicalMedicine 2024 Jul;73:102693 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102693] [Medline: 39429811]
76. de Zambotti M, Goldstein C, Cook J, Menghini L, Altini M, Cheng P, et al. State of the science and recommendations for using wearable technology in sleep and circadian research. Sleep 2024 Apr 12;47(4):zsad325. [doi: 10.1093/sleep/zsad325] [Medline: 38149978]
77. Lee T, Cho Y, Cha KS, Jung J, Cho J, Kim H, et al. Accuracy of 11 wearable, nearable, and airable consumer sleep trackers: prospective multicenter validation study. JMIR Mhealth Uhealth 2023 Nov 02;11:e50983 [FREE Full text] [doi: 10.2196/50983] [Medline: 37917155]
78. Banzi R, Camaioni P, Tettamanti M, Bertele' V, Lucca U. Older patients are still under-represented in clinical trials of Alzheimer's disease. Alzheimers Res Ther 2016 Aug 12;8(1):32 [FREE Full text] [doi: 10.1186/s13195-016-0201-2] [Medline: 27520290]
79. Gilmore-Bykovskyi AL, Jin Y, Gleason C, Flowers-Benton S, Block LM, Dilworth-Anderson P, et al. Recruitment and retention of underrepresented populations in Alzheimer's disease research: a systematic review. Alzheimers Dement (N Y) 2019 Jan;5(1):751-770 [FREE Full text] [doi: 10.1016/j.trci.2019.09.018] [Medline: 31921966]
80. Wilson SA, Byrne P, Rodgers SE. 'I'd be lost without my smartphone': a qualitative analysis of the use of smartphones and tablets by people living with dementia, mild cognitive impairment, and their caregivers. Aging Ment Health 2024 Apr 04;28(4):595-603 [FREE Full text] [doi: 10.1080/13607863.2023.2205585] [Medline: 37139944]

81. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2023 Mar 01;42(1):20-32. [doi: 10.1017/S0714980822000198] [Medline: 35912590]
82. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform 2018 Jan;77:120-132 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]

Abbreviations

AD: Alzheimer disease
EEG: electroencephalography
LBD: Lewy body disease
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
PD: Parkinson disease
PPI: patient and public involvement
PSG: polysomnography
PSQI: Pittsburgh Sleep Quality Index
RESTED: Remote Evaluation of Sleep to Enhance Understanding of Early Dementia
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by M O'Connell; submitted 19.02.25; peer-reviewed by N Misra, R Singh, R Tripathi; comments to author 11.03.25; revised version received 28.03.25; accepted 28.04.25; published 28.05.25.

Please cite as:
Gabb VG, Blackman J, Morrison H, Li H, Kendrick A, Turner N, Greenwood R, Biswas B, Heslegrave A, Coulthard E
Longitudinal Remote Sleep and Cognitive Research in Older Adults With Mild Cognitive Impairment and Dementia: Prospective Feasibility Cohort Study
JMIR Aging 2025;8:e72824
URL: https://aging.jmir.org/2025/1/e72824
doi:10.2196/72824
PMID:

Original Paper

Staff Enablement of the Tovertafel for Enrichment in Residential Aged Care: Field Study

Ryan M Kelly[1,2], BSc(Hons), MSc, PhD; Asmita Manchha[3], BCom, BBusMan(Hons), PhD; Jenny Waycott[2], BA(Hons), PhD; Rajna Ogrin[4], BSc, BPod(Hons), PhD; Judy A Lowthian[4,5,6], BAppSc(SpPath), MPH, PhD

[1]School of Computing Technologies, RMIT University, Melbourne, Australia

[2]School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia

[3]Silverchain, Melbourne, Australia

[4]Bolton Clarke, Melbourne, Australia

[5]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

[6]Faculty of Health and Behavioural Sciences, The University of Queensland, St Lucia, Australia

Corresponding Author:
Ryan M Kelly, BSc(Hons), MSc, PhD
School of Computing Technologies
RMIT University
124 La Trobe St
Melbourne, 3000
Australia
Phone: 61 (03) 9925 2000
Email: ryan.kelly@rmit.edu.au

Abstract

Background: Enrichment activities are essential for enhancing the psychosocial well-being of older adults living in residential aged care homes. There has been increasing interest in using digital technology for enrichment, but the implementation of technology requires careful support and enablement from staff to ensure that residents experience the intended benefits.

Objective: This study aimed to understand how care staff facilitate aged care residents' use of the Tovertafel ("magic table" in Dutch), a technology that projects images onto a tabletop to enable groups of people to play games. The study further aimed to understand the benefits arising from the Tovertafel when facilitated by staff.

Methods: We conducted a field study in 1 residential aged care home in Queensland, Australia. The methods included semistructured interviews with the staff and residents about their experiences with the Tovertafel, observations of 4 sessions in which the residents and staff played Tovertafel games, and a diary completed by the staff after Tovertafel sessions. Data were analyzed through reflexive thematic analysis.

Results: We developed 3 themes through our analysis. Theme 1 highlights the need for the staff to overcome physical and personal barriers before Tovertafel sessions could take place. These included a lack of a dedicated space for playing Tovertafel games and the residents' reluctance to attend Tovertafel sessions. Theme 2 highlights how the staff used creative strategies to make Tovertafel sessions successful. These included helping the residents learn how to interact with the games; adapting the activity to suit the capabilities of the residents; sustaining engagement by choosing appropriate games; and using prompts, questions, and storytelling to make the games more engaging. Theme 3 describes the benefits and outcomes that arose from staff-supported enablement of the Tovertafel, including participation in an enjoyable physical activity, socialization, and reminiscence.

Conclusions: This study suggests that the Tovertafel provides opportunities for aged care staff to engage in creative play and personalization catering to residents with different capabilities. However, the benefits arising from the Tovertafel are unlikely to be achieved without substantial facilitation from the staff, who play a key role in enabling the participation of the residents. Sustaining the engagement of the residents is important during Tovertafel activities and can lead to beneficial outcomes.

(JMIR Aging 2025;8:e67919) doi:10.2196/67919

KEYWORDS

aged care; care staff; creative care; older adults; play; residential aged care; social enrichment; technology; Tovertafel

Introduction

Background

In 2023, a total of 185,000 Australians aged ≥65 years lived in residential aged care homes [1]. Older individuals usually move to residential care when they require high levels of support to enable appropriate clinical care alongside holistic well-being [2,3]. Holistic well-being refers to addressing residents' physical, psychological, and social needs [4]. To promote physical well-being, staff members from lifestyle teams support residents with activities of daily living, such as bathing, dressing, and eating, and enable them to participate in physical activities through exercise programs.

To promote psychological and social well-being, staff members provide an array of enrichment activities, such as art, gardening, music, games, and excursions. These activities are important for encouraging social interaction and preventing boredom [5], enhancing residents' quality of life [6], and promoting a sense of continuity, particularly when residents engage in activities that interest them [7]. However, activities in residential care are often group based and may not cater to the needs of all residents [6], particularly when there is diversity among residents regarding individual interests [7], cultural background [8], or functional and cognitive capabilities [9,10].

Recent years have seen an increasing uptake of digital technologies for enrichment in residential care homes [11]. Technology-based activities can both expand the range of available options and enable new experiences of joy and social engagement [12]. For example, immersive virtual reality (VR) offers the possibility of virtually "traveling" to distant locations, revisiting places from the past, and remotely attending concerts [13,14]. When used appropriately and in a way that demonstrates person-centered care, technology-based activities can address the individual needs and goals of residents, including those with cognitive impairment [11].

However, research has identified that appropriately trained staff are often required to enable the effective use of technology in aged care [15]. Residential care staff are known to be dealing with high workloads [16], and there is a risk that facilitating technology may add to this load. While technology facilitation can be enriching when it involves meaningful care work [17], significant work in setting up and manipulating the technology can be an unwanted burden that limits the use of the technology [18,19]. Therefore, evaluating the demands placed on care staff, along with the potential benefits of using technology with residents, is an important consideration when justifying the use of technology as part of a care home's lifestyle and social calendar.

The Tovertafel

This study focuses on staff involvement in enabling the Tovertafel ("magic table" in Dutch), a commercial system that enables groups of people to play games by tapping on moving images projected onto a tabletop [20]. The games vary in difficulty, from simple activities such as sweeping leaves or popping bubbles to more advanced puzzles and team-based games that require cognitive and social engagement. While the Tovertafel is aimed at providing stimulation for care home residents living with dementia, it can be used by residents with various levels of cognitive impairment. Table 1 provides examples of Tovertafel games.

Table 1. Examples of Tovertafel games.

Game name	Description
Leaves	Players use arm movements to sweep piles of leaves from the tabletop.
Football	Players try to score goals at either end of the table by tapping a soccer ball.
Music Box	Players tap on musical notes rotating around a music box in the center of the table. The notes play piano sounds when tapped. At the end of the game, the music box reveals a spinning ballerina and plays a song.
Rhymes	Players tap on feathers to reveal the words of a nursery rhyme, which can then be read aloud.
Sayings	Players tap on moving bubbles to reveal a phrase or saying, for example, "Never look a gift horse in the mouth."
Silverware	Players wave their arms over images of cutlery to simulate polishing a dinner set. Once all pieces are polished, plates and food appear on the tabletop, simulating the table being laid at dinner time.
The Veggie Patch	The game simulates gardening by requiring players to plant and water vegetables. Seed packets move over the table, causing vegetables to be planted when tapped. There are also moving clouds and sunbeams, which cause the plants to grow when tapped.

A small number of studies have examined the use of the Tovertafel in residential aged care. These studies have primarily focused on identifying benefits and drawbacks of the technology. Good et al [21] interviewed care staff about the perceived benefits of the Tovertafel for residents living with moderate to advanced dementia. Staff felt that using the Tovertafel had several benefits, including improved communication between staff and residents and short-term improvements in residents' mood. However, the Tovertafel was rarely used by residents without initiation from carers or loved ones, and residents with dementia could only use the Tovertafel for a short time because of impaired concentration. A study by Talman and Gustafsson [22] identified benefits of the Tovertafel for people living with intellectual disabilities. However, these findings may not be applicable to aged care. The Tovertafel's official website lists various studies that have been conducted to validate the system, but most are unpublished, originate from master's theses, or require readers to contact the company to obtain further information [23]. Other than the studies mentioned earlier

[21,22], the only publicly available study evaluating the Tovertafel is in Dutch [24].

While the impact of the Tovertafel may be positive, there has been limited examination of the care strategies required to facilitate the participation of residents in the activity. There is also limited understanding of how staff work to overcome issues that may impact the engagement of aged care residents with the Tovertafel. Research is needed to better understand how staff facilitate the Tovertafel activity and to inform decisions about using this technology, given the limited resources available in many residential aged care homes.

Objectives

This study aimed to address the following research questions:

1. How do care staff enable residents' use of the Tovertafel in residential aged care homes?
2. What benefits arise from staff-supported enablement of the Tovertafel?

Methods

Ethical Considerations

All procedures in this study received approval from the University of Melbourne's Human Research Ethics Committee (ID 12900) and the Bolton Clarke Human Research Ethics Committee (approval 230001). Our study involved participants who were either staff or residents from one aged care home in Australia. The staff participants signed a participant information sheet and consent form (PICF) and received an Aus $20 (US $13.37) gift voucher as a token of appreciation.

The care home residents who participated in the interviews were provided with a PICF that included pictures of Tovertafel games, along with a textual description of the technology. Some of these residents were shown a YouTube video of the Tovertafel if they were unable to understand its functionality from the PICF alone. The residents who agreed to be interviewed provided informed consent by signing the PICF and received baked goods as a token of gratitude. Some residents living in the memory support unit of the care home did not have the capacity to consent, so interviews were not sought about their experiences with the Tovertafel. However, we observed the sessions in which these residents participated and collected observational notes about their use of the Tovertafel, without capturing personal or identifying information, in line with the ethics approvals for the project.

All data from the study was storied securely in password-protected filing cabinets or secure cloud-based storage to protect the confidentiality of the data. All transcripts were anonymized, and names were replaced with pseudonyms to protect participants' privacy.

Study Design

We conducted a field study to investigate how the Tovertafel was used in 1 Australian residential aged care home. The study involved semistructured interviews, observations, and a participant diary. We chose these methods to gain in-depth insight into staff members' experiences of facilitating the Tovertafel and to understand how this facilitation elicits benefits for residents. All procedures were developed by the research team, drawing on our previous experience with the chosen methods [25,26] and our experience of conducting fieldwork in aged care [8,27,28].

First, the staff who were familiar with using the Tovertafel were interviewed about their experiences of using the system, the strategies that they had used to make the activity successful, and the perceived benefits. We also conducted interviews with the residents who had used the Tovertafel, provided they had the capacity to give informed consent and provide feedback. These interviews sought to understand what the residents liked about the activity and what they found challenging. The interview schedule for the residents contained simple questions to avoid overburdening them.

Second, we observed 4 activity sessions, each lasting 45 minutes, in which small groups of staff and residents played a variety of Tovertafel games. These sessions were scheduled by the lifestyle team of the care home as part of the usual activity schedule. The observations enabled us to understand how the Tovertafel was used in practice, what the staff did to make each session successful, and how the residents responded to the activity.

Third, the staff leading the Tovertafel activity were asked to complete a diary after each session. The diaries invited further reflection on the use of the Tovertafel, including the views of the staff about the impact of the Tovertafel on the mood of residents, what worked well, and what the staff found challenging about the session (refer to Multimedia Appendix 1 for the diary design). The staff were asked to complete diary entries after each Tovertafel session to capture their experiences as close to the time of delivery as possible.

Table 2 lists the observed Tovertafel sessions, the number of people involved, and the timing of diary entries completed by the staff members.

Table 2. List of the Tovertafel sessions examined.

ID	Date and time	Location	Number of people involved	Observed by researchers	Diary entry completed by staff
1	June 6, 2023, Tuesday, 9:15 AM-10 AM	The memory support unit	2 residents, 2 lifestyle staff, and 1 nursing staff	✓	
2	June 7, 2023, Wednesday, 9:15 AM-10 AM	The memory support unit	4 residents and 2 lifestyle staff	✓	✓
3	June 12, 2023, Monday, 9:15 AM-10 AM	The memory support unit	7 residents, 2 lifestyle staff, and 1 nursing staff		✓
4	June 21, 2023, Wednesday, 1:30 PM-2:15 PM	The nursing home	4 residents and 1 lifestyle staff		✓
5	June 27, 2023, Tuesday, 9:15 AM-10 AM	The nursing home	3 residents and 2 lifestyle staff	✓	✓
6	June 28, 2023, Wednesday, 9:15 AM-10 AM	The nursing home	6 residents and 2 lifestyle staff	✓	✓

Study Site

Our study took place at 1 residential aged care home of a not-for-profit aged care provider in Queensland, Australia, in 2023. The home provides care for approximately 80 residents divided across 3 residential wings: a memory support unit for residents with severe cognitive impairment, a nursing home area for other residents with high care needs, and a hostel area for residents with lower care needs. The home had 112 staff at the time of our study, including, for example, administrative staff, nurses, managers, and housekeeping staff. The home had 2 staff members who oversaw lifestyle activities: 1 diversional therapist and 1 lifestyle activities coordinator. These staff members were supported to run activity sessions by nursing and personal care staff, depending on availability and the number of people required.

The residents at the home predominantly follow a daily routine with fixed times for meals and personal care. Enrichment activities are scheduled around this routine. The lifestyle activities coordinator prepares a monthly activities calendar, which is distributed to residents on paper. Each day has an activity in the morning and another in the afternoon. Activities last between 45 and 60 minutes, and the residents are free to attend as they please. All activities are announced by staff over a public address system transmitted across the 3 residential wings.

The care home has a large room for providing lifestyle activities. At the time of our study, this room was undergoing renovation. Because of this, the equipment for the Tovertafel had been installed in the dining areas of the memory support unit and the nursing home.

During this study, the Tovertafel activity was scheduled in the activities calendar, allowing the residents from the memory support unit and the nursing home to participate (Table 2). The residents from the hostel area were also welcome to participate. The staff had been using the Tovertafel for 2 years, having received the device as a gift from a benefactor in 2021, but it had not been included in the activities schedule for >6 months before this research. The staff had not received formal training on how to use the Tovertafel. Instead, they learned on the job. After the Tovertafel was first installed in the home, the diversional therapist downloaded a PDF of the system's user manual to understand how the device should be operated.

Participant Recruitment

The staff members were invited to participate in the study if they were familiar with the Tovertafel or if they were scheduled to help facilitate the sessions during the study period. All 6 staff members who met either of these criteria agreed to participate. Volunteers and family members were also eligible to participate in the study, but none were present when data were collected.

Procedure

All procedures were carried out by 2 researchers (RMK and AM). Two fieldwork visits, each lasting 4 days, were conducted in June 2023. Both researchers spent several days on-site before and after the Tovertafel sessions, providing opportunities to learn about the social routines of the care home [29]. The researchers took photographs of the setting, without capturing images of the residents or the staff, and tested the Tovertafel games to understand what they involved. The dates and times of the observed sessions are provided in Table 2. The researchers created private, written fieldnotes after each day of the study to record their experiences and personal reflections about the care home.

During the first fieldwork visit, the researchers interviewed the staff members about their experiences with the Tovertafel and observed 2 Tovertafel sessions in the memory support unit. The researchers conducted nonintrusive observations of each session, observing the staffs' and the residents' interactions around the Tovertafel and with each other, from 2 to 3 meters away. The researchers took notes about which games were played in each session, what the staff did to facilitate the activity, and how the residents responded to the Tovertafel, including emotional responses and social interactions. All observations of emotional responses were noted in a free-text form, rather than using a predefined behavioral grid. The staff member in charge of each session was requested to fill out the research diary after they had facilitated the activity. All diary entries (5 in total) were completed by the coordinator of the lifestyle activities in the

care home, who attended all the Tovertafel sessions examined during the study.

The first fieldwork visit was followed by a 2-week gap during which the researchers were off-site. Two Tovertafel sessions were scheduled by the staff during this time (Table 2). These sessions were not observed by the researchers, but the lifestyle activities coordinator completed the diary after each one.

During the second fieldwork visit, the researchers conducted observations of 2 Tovertafel sessions in the nursing home and invited the residents to provide feedback through interviews after each session. Follow-up interviews were conducted with the staff who had facilitated the sessions.

Analysis

Data consisted of interview transcripts, observational and fieldwork notes, photographs, and diary entries. Interviews were transcribed using a professional transcription service. Data were analyzed using the 6 phases of the approach to reflexive thematic analysis by Braun and Clarke [30]. Reflexive thematic analysis is interpretivist and acknowledges that the subjectivity of the researcher can be a valued analytical resource, while also recognizing that the analysis is influenced by the positionality of the researcher [31]. This analysis was led by RMK, whose position is that technology has the potential to enrich the lives of people living in residential care but risks becoming an unnecessary burden, particularly if the technology is not well designed for this setting [17,26]. This position informed the analysis by highlighting both the burdens involved in enabling the Tovertafel and the enriching aspects of the work, particularly when it provided the staff with opportunities to enhance the activity for the residents.

Following the 6-phase approach by Braun and Clarke [30], RMK first read through the data before inductively creating semantic and latent codes at the sentence level [31]. RMK collated the codes into potential themes that captured interpretations of the problems the staff had faced, the ways that they facilitated sessions, and the perceived benefits of the Tovertafel. The analysis was discussed several times with AM, who iterated the themes based on her involvement in the data collection. RMK then drafted the themes into an initial version of the paper, which was discussed and reviewed by all authors [32]. This team input supported a critical and balanced perspective on the Tovertafel activity. After receiving feedback from peer reviewers, the themes were revised and reorganized to strengthen the interpretations.

In the following section, we present 3 themes. These are organized chronologically, as the themes we developed align with a before-during-after framing of the activity and reflect the temporal ordering of actions conducted by the staff when facilitating the Tovertafel sessions. The first theme highlights the need for the staff to overcome physical and individual barriers *before* initiating the Tovertafel activity. The second theme reveals the importance of creative care from the staff *during* Tovertafel sessions. The third theme details the benefits and outcomes *arising from* staff-supported facilitation of the Tovertafel. Verbatim quotes are used to illustrate findings, with pseudonyms and roles at the care home used to identify participants. Descriptions of the residents' emotions or reactions to the activity are based on observations by the researchers or entries made in the staff diaries.

Results

Participant Characteristics

All 6 staff members interviewed were women. The interviews lasted between 15 and 38 (mean 30, SD 8.7) minutes. Table 3 lists participants' pseudonyms, roles, and time at the workplace for the staff participants. Each Tovertafel session was led by Corina, the lifestyle activities coordinator, along with at least 1 other staff member.

Table 3. List of interviewees.

Pseudonym	Gender	Role	Time at the workplace	Time in current role	Interviewed during the first fieldwork visit	Interviewed during the second fieldwork visit
Anne	Woman	Care home manager	20 y	4 y	✓	
Brenda	Woman	Diversional therapist	11 y	4 y	✓	✓
Corina	Woman	Lifestyle activities coordinator	4 mo	4 mo	✓	✓
Diane	Woman	Special care unit nurse	16 y	16 y	✓	
Ellen	Woman	Personal care worker	11 y	11 y	✓	
Fiona	Woman	Enrolled nurse	8 y	3 y	✓	
George	Man	Nursing home resident	N/A[a]	N/A		✓
Herbert	Man	Nursing home resident	N/A	N/A		✓
Ingrid, Julie, and Karen	Women	Hostel residents (interviewed together)	N/A	N/A		✓

[a]Not applicable.

In addition, 5 residents (2 men and 3 women) were interviewed after the Tovertafel sessions. These interviews lasted between 10 and 19 (mean 14.3, SD 4.5) minutes. Other residents were approached for the interview but either declined to provide feedback, were unable to participate because of severe communication impairments, or were occupied with personal care activities.

Theme 1: Initiating Tovertafel Sessions Required Attending to Physical and Personal Barriers

Summary

Our analysis showed that the staff needed to overcome multiple barriers before running the Tovertafel activity. These included the lack of a dedicated space for play and the need to address residents' uncertainty about joining the activity. These barriers reflect the physical and personal challenges associated with using group-based digital technology, such as the Tovertafel, within residential care homes. We provide further explanations of physical space and perceptions of technology as subthemes in the following sections.

Physical Space: The Care Home Lacked a Dedicated Space for Play

At the time of our study, the care home lacked a dedicated activities space because of an ongoing renovation in the activities room. All group activities had to take place in other communal areas.

Equipment for the Tovertafel system, including an electrical connection and a ceiling bracket, had been installed in the dining areas of the memory support unit and the nursing home. These areas were typically set up with tables and chairs in preparation for mealtimes. To make each space suitable for the Tovertafel activity, the staff had to reconfigure the furniture to create a larger table suitable for the residents to sit around.

The staff did not feel that this practical work was especially burdensome, although it did require additional effort. As part of the setup, the staff also placed a white cloth over the table for use as a playing surface, after noting that the games were hard for some residents to see without using the tablecloth. This was because the wooden tabletops had not been designed for use with the Tovertafel and tended to absorb the projected images, making a lighter surface necessary. In addition, the dining areas had large windows that let in considerable sunlight, making it harder for residents to see the projections if the tablecloth was not used:

> *The setup doesn't take that long...but one thing I found was that the graphics were a bit too light. So maybe it needs to be a certain colour that it projects onto. Because I noticed that with a few of the games, [the residents] did struggle to see what it was.* [Diane, care nurse]

The residents noted that they were hesitant about attending Tovertafel sessions because of the lack of a dedicated space for play. As noted earlier, the activities room at the care home was undergoing renovation. The residents from the hostel area where the residents had lower care needs—were invited to join the activities that took place in the nursing home and memory care unit. However, they expressed reluctance to enter these areas and felt that that it would be better if the Tovertafel was in a shared activities room:

> *I was wondering if it's possible to have that in the activity room. More people might go then...a lot of people down here don't like going into the nursing home.* [Julie, hostel resident]

In addition to needing to overcome physical space barriers to initiate the Tovertafel sessions, the staff also had to ensure that they were able to reset the space at the end of each session. We observed that there was a risk of conflict if the chairs and tables were not rearranged in the same order as before. This was because specific residents could not be seated together, and the staff needed to place the tables carefully to minimize extra work for the care and culinary staff during mealtimes. This situation also meant that the Tovertafel could not be left available for the residents to use at their convenience. As highlighted in previous work [21], the staff thought that the system might be used more often if residents could use it with visiting family members, without the need for staff assistance.

Perceptions About Technology: Residents Were Often Reluctant to Join Tovertafel Sessions

A second barrier that the staff had to overcome was the reluctance of the residents to join the sessions. The staff found it challenging to encourage the residents to attend the Tovertafel sessions because many were not interested in technology-based activities:

> *I find they're not really interested in technology as a lot of the residents have farming backgrounds. They were used to milking the cows and doing all the farm work.* [Ellen, personal care worker]

> *People tend to go "under cover" when you mention the word technology.* [Corina, lifestyle activities coordinator]

In addition, the residents were often uncertain about what the Tovertafel involved if they had not seen it before. Most of the residents we encountered spoke English as their first language, but the word "Tovertafel" lacks obvious meaning for those who do not speak Dutch. The staff found that describing the Tovertafel as "the table game" or "something you play with your hands" was more effective, though still challenging without demonstrating the activity in person.

Because of this, the staff sought to increase attendance by going to the rooms of the residents before sessions took place to explain what the activity would involve and to assure the residents that it would be enjoyable. The staff also approached the residents who were in the nearby common areas when the activity began and asked them if they would like to give the activity a try. The staff felt this was different from other frequently facilitated activities (eg, choir and concerts), which the residents did not need encouragement to attend.

The issue of attendance was important because the staff felt that the Tovertafel worked best with groups of 4 to 6 players. Having 4 players was seen as too few to make the games enjoyable. This meant that achieving enough players was important for

the activity to be successful. The staff told us that the Tovertafel had not been used for some time before the study because attendance had been poor:

> *We have used it before, but we reached a stage where we weren't getting a lot of interest. So, I've taken it off the activity calendar for a time. We just weren't having the numbers turn up to the activity.* [Brenda, diversional therapist]

The staff were also conscious that running the activity too often with the same people risked creating disinterest. This view was echoed by the residents:

> *I think if we played it every day, it might get boring. Once a week, that's no problem.* [George, nursing home resident]

To initiate the Tovertafel activity, the staff needed to be aware of which residents were available to join, how many were willing to participate, and had to encourage those who were available to attend. Addressing these social dynamics was challenging and time consuming but was essential to ensure that sessions had sufficient participants to make the games worth playing.

Theme 2: The Staff Used Creative Care Strategies to Guide the Tovertafel Sessions

Summary

This theme highlights how the staff creatively enabled the use of the Tovertafel *during* the activity; that is, once the staff had created the conditions for a Tovertafel session by arranging the play area and encouraging the residents to attend, additional work was needed to make the sessions successful. We observed that the staff used creative approaches to help residents learn the activity, enable the continued participation of the residents, and sustain their engagement. Each of these activities required differing forms of creativity, which are described in the subsequent sections.

Selecting Gentle Games to Enable Initial Learning

At the start of each session, we observed that the residents were typically unsure what the Tovertafel was and what it involved. In 1 diary entry, Corina, the activities coordinator, similarly noted that the residents were hesitant at the beginning of a Tovertafel session but became more confident once they understood what to do:

> *The mood was a bit daunting and [residents were] unsure what was going on but once we got started the mood changed and [they] became interested.* [Corina, lifestyle activities coordinator, diary excerpt, session 2]

This highlighted the need for the staff to help the residents learn how to use the Tovertafel. The staff had identified a creative way to achieve this learning. At the start of each session, they chose to leave a simple game running on the tabletop while helping the residents join the activity. The staff typically chose the game Leaves for this situation. This game involves sweeping leaves from the tabletop using arm movements. We observed that playing with the leaves helped the residents to understand what would happen if they touched the images. Changes in the leaves were both noticeable and easy to trigger, helping the residents understand how the Tovertafel reacts to their movements.

In this way, starting each session with a gentle game provided opportunities for social learning, whereby the residents learned about the activity through watching others. In session 2, we observed a resident who repeatedly exclaimed that she could not join in because she "didn't know what it was all about." She eventually began playing after watching the interactions of other residents and after receiving encouragement from staff.

These instances highlight that creative work was needed to make the residents feel comfortable at the outset of each session. They also reveal that the staff were not simply supervising the activity. Rather, they were central to making the residents feel at ease and actively encouraged the participation of the residents during the sessions.

Developing Creative Solutions to Ensure Safe Participation

Creativity was evident in the way that the staff developed solutions to other problems that hindered the participation of the residents. For example, some residents found it challenging to trigger a response from the Tovertafel because of mobility impairment. Many could not stand up without assistance and therefore could not reach across the table, while others had difficulty moving their hands or arms.

Recognizing this, the staff had cut up a "pool noodle" (ie, a cylindrical flotation device made of foam) into pieces, which the residents then held to extend their reach and trigger the Tovertafel more easily:

> *Some [residents] have problems with their arms, or they've hurt their elbow in some way. So having the pool noodle, they don't have to reach. And a lot of residents tire very quickly when they're using the arms.* [Brenda, diversional therapist]

The staff were conscious that some residents might experience motion sickness when using the Tovertafel. In session 2, the staff paid close attention to a resident who was known to experience vertigo in response to moving images. The staff described the importance of attending to visual cues, such as body language and facial expression, while running the game to ensure that the residents remained safe. In this way, the staff exhibited creative care by monitoring for signs of distress during the games and changing to a simpler game if required:

> *You need to go by the visual cues and that's probably the advantage of Tovertafel. It's a visual reaction. And it can depend day to day on that resident, with how they're going to interact in an activity or with the Tovertafel.* [Anne, care home manager]

Sustaining Engagement by Choosing Appropriate Games

The importance of choosing games was mentioned repeatedly by the staff during the interviews. We observed that the staff always took responsibility for selecting games, using a remote control to pick games from the Tovertafel's menu system.

The staff explained that choosing appropriate games was essential to sustain the engagement of the residents. One reason was that the staff felt that there needed to be a good match between the games on offer and the physical or cognitive ability of the residents in attendance:

> *[It] depends on who you've got present in the session, because they've all got varying attention spans.* [Anne, care home manager]

In addition, the games needed to be relevant to the interests of the residents; otherwise, the sessions risked becoming boring. The staff explained how this risk could be mitigated by choosing games with features that appealed to the residents. Some Tovertafel games replicate tasks such as gardening and cooking. These games were perceived as engaging because they are modeled on familiar leisure activities. For instance, the staff often selected the Veggie Patch game because many residents were interested in gardening:

> *The gardening game is one where many of the ladies enjoy gardening, so we will do that and ask what they used to do in their garden. There's one person who got very engaged by that. She loves gardening and that's one way to try and get her to engage in the game because gardening was her hobby.* [Brenda, diversional therapist]

Sometimes the staff chose games that were known to be exciting and avoided games that they thought were boring. This was established through trial and error, as well as through verbal and nonverbal feedback from the residents. The staff also switched between games based on the current state of the residents. For example, if residents appeared to be losing interest in a gentle game, staff would switch to one involving physical activity to help engage them:

> *I find if they play it for too long, they get a bit tired of it and they get a little bit bored. So that's why you've got to change to a different game, just to change the mood. It seems like they like something a bit quicker as well, more so than something that's a little bit slow and repeated.* [Corina, lifestyle activities coordinator]

Finally, the staff had identified that staying with the same game for too long risked creating boredom among the players. Changing Tovertafel games during a session was essential to counter this risk, and the staff typically spent no longer than 3 to 5 minutes on each game. Tovertafel games have various difficulty levels, and some were seen as too simple to be played for more than a few minutes. The staff tried to introduce variety by including different types of games within a session, either by manually choosing games or using the Tovertafel's built-in shuffle function, which moves through games in a randomized order:

> *I don't stay on one game for too long, like maybe five minutes. Or I gauge by watching the people around the table and if they're disengaging or not finding it interesting enough, I'm like, okay, let's switch another game now. You know, to try and give them a bit of variety. And to keep their focus.* [Brenda, diversional therapist]

Using Prompts, Questions, and Storytelling to Make the Games More Engaging

A final creative strategy was the use of different spoken techniques to keep the residents engaged. Sometimes this involved gently prompting the residents. For example, a resident in session 4 was observed to be falling asleep during the game. The staff offered words of encouragement—"Come on George," "Wake up George"—to keep him interested. In session 5, a resident felt that she "wasn't doing it right" and stopped interacting with the game. A staff member provided gentle reassurance, saying "No, you're doing a good job," which encouraged her to become involved again.

Beyond this gentle prompting and encouragement, the staff tried to keep the residents' attention on the activity by creatively responding to elements of the games with questions that encouraged interaction. For example, during the Rhymes game, the staff chose to read each rhyme aloud and then asked the residents questions about them. These questions were not part of the game but were added by the staff to make the activity more engaging. In 1 instance, after seeing a rhyme from the song "Rock-a-bye baby," the staff asked the residents if they knew the rhyme and encouraged them to sing along. During the interviews, the staff explained that this was another technique to maintain the residents' focus:

> *To make it work, I keep communicating with the residents and try not to get them off track. Like, you know, just make them stay involved. Otherwise, they're going to drift off and get sidetracked with the game.* [Corina, lifestyle activities coordinator]

Other games involved solving puzzles to reveal pictures of places or animals. The staff used these as opportunities to invite conversation about the past, asking the residents questions such as, "Have you ever been to New Zealand?" or "Did you used to have chickens in your garden?" These questions prompted the residents to talk about their lives. Staff members noted that specific games provided inspiration for these conversations, leading to positive outcomes:

> *What I try to do is to get them to share their stories, so for example when we play the music box game, I might ask them if they had a music box when they were younger. The game with silver polishing is something they would have done when they were younger. You can ask them about that, ask who they had [over] for dinner and what they used to eat.* [Brenda, diversional therapist]

These behaviors were creative techniques that had been instigated by staff members. They were not a standard part of the Tovertafel games, nor were they part of any training that the staff had received. Rather, they were strategies that staff had found to work well and represented the attempts of staff members to build on the basic elements of the Tovertafel games. This appeared to make the activity more active and engaging, as opposed to one in which the residents passively interacted with the system without understanding what they were doing.

Theme 3: Staff Creativity Fostered Benefits and Outcomes

This theme highlights 3 benefits arising from the staff-supported enablement of the Tovertafel. These included moments of joy and fun from participating in a physical activity, opportunities for socialization with other residents, and reminiscence prompted by elements of the games.

Participating in a Fun Physical Activity

The staff universally viewed the Tovertafel as a positive activity for the residents, and in their view, the benefits outweighed the work involved in setting up and running the technology.

The staff described how keeping the residents engaged was important for enjoyment, otherwise attendees would receive limited value from the activity. Through supporting the residents to use the Tovertafel, we observed that the staff were able to elicit moments of joy and laughter, particularly when playing games that involved considerable physical activity. For example, in 1 diary entry, Brenda noted a positive outcome from the session:

> *Happy, smiling during the Tovertafel, lots of laughter. There was a change, they were feeling unsure but once participating their mood became more alert and wanted to give it a go.* [Brenda, diversional therapist, diary excerpt, session 3]

A related benefit was that the Tovertafel encouraged physical activity through the use of the hands and arms. The staff viewed this as a rare opportunity for the residents to engage in gentle exercise, as illustrated by the following comment:

> *There is a whack-a-mole game, and I like that because we've got people with mobility issues and so it's good exercise for them.* [Fiona, enrolled nurse]

Encouraging Social Interaction Between Residents

The Tovertafel provided an opportunity for the residents to socialize with others in the care home. We observed that, outside of mealtimes or scheduled activities, many of the residents would either be watching television or spending time alone in their room. The residents commented on how they enjoyed the social aspects of the activity:

> *I think it's more of a social occasion, when you play games or whatever you play, you can get to know people that way and everybody has different characters, you know what I mean?* [George, care home resident]

The staff felt that the Tovertafel enabled socialization in a unique way by encouraging residents to interact with people they might not usually engage with:

> *It gets them together, you know, like having a chat. So, there's just that bit more interaction. I think it's way better than the TV. And compared to our trivia games, not everyone gets involved in the trivia. Whereas with [Tovertafel] you can involve most of them. And even if they don't want to play it, at least they might just come and watch. They're still interacting in some way, you know?* [Diane, special care unit nurse]

As Diane noted, some residents did not always want to play with the Tovertafel, preferring to sit and watch. Others attended the sessions but were unable to take part because of physical impairment. The staff believed that the Tovertafel still offered a positive opportunity for these residents to socialize, either by watching the activity or by taking on a supporting role in the games. One staff member explained as follows:

> *The ones who honestly can't play, I think it's a good idea, even if they are there to join in and give them something to do, just remember a score or help out because then they are participating and joining in as well. So that's a good thing. You're not inviting them to sit there and fall asleep. The whole idea of it is to keep them inspired, just to let them know that we're here to have a good time and they're in a social environment.* [Corina, lifestyle activities coordinator]

Fostering Shared Reminiscence

Finally, the Tovertafel games provided opportunities for shared reminiscence. However, this reminiscence only occurred when the staff prompted the residents to share their stories in response to the games. By asking questions based on the game elements, the staff actively tried to prompt memories about the past. This included topics such as whether the residents used to keep animals in their garden, what they used to make for dinner, and what they typically did to celebrate birthdays—all of which were embodied within different games. This was seen as a particularly positive outcome of using the Tovertafel, as explained by 1 staff member:

> *Going back in time is good for them to feel good within themselves because it brings back memories. Like this morning, there were some songs that the residents remembered from their younger years. To me that's a good thing, reminiscing.* [Corina, lifestyle activities coordinator]

Prompting reminiscence enabled the residents to share their stories and learn about one another. In session 2, we observed that this reminiscence led to a protracted conversation in which a resident shared a story about a country property she helped renovate. The resident later retrieved a photo album from her room and began showing pictures of the renovation works to other residents. Interactions like these enabled the staff to engage in "biography work" [33], which involves learning about residents so that staff can address their needs using more culturally responsive practices. In this way, reminiscence appeared to be beneficial for the residents and the staff members attending each session.

Discussion

Principal Findings

Overview

This study aimed to understand how the care staff enable the Tovertafel in residential aged care and how this enablement contributes to benefits for residents. Our findings suggest that

staff facilitation was crucial for the effective use of the Tovertafel. This began with overcoming barriers to facilitating sessions, including the lack of a dedicated space for play and the need to encourage the residents to become involved. The staff then played a central role in supporting the residents through the activity, using creative techniques to guide the residents through the games and sustain engagement, while paying close attention to the safety of the residents. These actions helped residents gain meaningful benefits from the Tovertafel, which included fun physical activity, socializing with others, and opportunities for shared reminiscence.

Staff Enablement of the Tovertafel

Our findings highlight the central role the staff played in enabling the Tovertafel and the importance of creative care strategies in sustaining the participation of the residents. Previous work has argued that creativity can be required when providing enriching technology-mediated activities in care homes [34]. Our findings support this claim by showing how the care staff engaged in creative work during Tovertafel sessions. The Tovertafel provides a range of games, many of which are very simple. Although it may be hoped that the residents can play these games without staff oversight, our findings show that facilitation by the staff was needed to ensure that the residents understood the purpose of the activity, how to play the games, and how to interact with the system. In addition, the staff undertook a variety of creative actions to enhance the engagement of the residents *during* the sessions. The staff were conscious that the residents might choose to disengage from the Tovertafel, given the simplicity and repetitive nature of the games and because some of the residents had limited attention span. Creative actions around the selection and duration of the games helped to mitigate these risks and kept the residents' focus on the activity. These creative actions, in turn, were likely essential to fostering the benefits and outcomes observed in this study.

However, the involvement of the care staff was also needed to instigate the sessions. This work is arguably less desirable than tasks involving creative care. We observed how staff time was needed to set up the physical space for play—partly a result of where the Tovertafel was placed in the care home. The staff and the residents felt that the activity might be more popular if it were available to use in a dedicated space, such as an activities room. However, previous work has shown that when the Tovertafel was set up in such a space, it was still rarely used by the residents without prompting from caregivers [21]. This implies that facilitation may still be required *during* the use of the Tovertafel, irrespective of where it is located. We also found that staff time was needed to encourage the residents to attend the sessions, given residents' lack of experience with technology and uncertainty about what to do. These efforts can be viewed as a burden, but they also provided an opportunity for the staff to engage directly with the residents and understand their reasons for nonparticipation. The staff were then able to factor this into the running of the sessions, such as by starting with a simple game to help residents to become comfortable with the technology.

Finally, the care staff were involved in making sure that the residents remained safe during the activity. In residential aged care, many residents are frail or are living with cognitive impairment. This was true for many of the people who participated in the sessions we observed. The care staff ensured that these residents remained safe while using the Tovertafel. The staff believed that the Tovertafel involved some risks for the residents, particularly those who experience dizziness, and so they were vigilant about resident safety. Previous studies have highlighted the need for staff to ensure the safety of residents when using "risky" technologies such as exergames [35] or immersive VR [15,27]. Our study shows that the involvement of the staff was similarly essential to ensure that the residents remained safe when interacting with the Tovertafel.

Benefits of Using the Tovertafel

Our findings shed light on benefits arising from the staff-supported enablement of the Tovertafel. First, Tovertafel encouraged physical activity and provoked moments of fun and joy for the residents. It can be difficult to provide stimulating physical activities in residential aged care because of high rates of impairment among the residents [36]. Our observations suggest that the Tovertafel promotes gentle upper-body exercise but can require creative problem-solving and support from the staff to make the activities accessible. We observed this when the staff triggered interactions on behalf of the residents and through their efforts to make the game easier to play by using pieces of foam to extend the reach of the residents.

A second benefit was that the Tovertafel provided an opportunity for the residents to socialize with each other and with the care staff. Although care homes such as the one we studied frequently provide a range of lifestyle activities, opportunities for social interaction can remain scarce [37], contributing to experiences of loneliness [38]. To address these issues, previous work has explored the benefits of technologies that encourage "ludic engagement" through playful interactions between the residents [39,40]. The staff in our study felt that the Tovertafel was similarly beneficial because it brought residents together for playful, face-to-face interaction. This may be hard to achieve with technologies that are single-user focused, such as VR [15]. In addition, the staff noted that the Tovertafel demands cognitive engagement and encourages conversation. This was perceived to be more beneficial than passive activities, such as watching television. It was also perceived as beneficial for the residents who participated as "active audience members" [41], observing while others engaged with the Tovertafel.

A third benefit was that the Tovertafel enabled the staff to prompt reminiscence among the residents. When used effectively, reminiscence can be a positive activity for older people living in residential care [42] and can be supported using technology such as VR [43,44] and digital mapping systems [45]. However, these efforts are also typically single-user focused, and it has been found that while care home residents often engage in reminiscence with family and friends, they do so less frequently with other residents and the care staff [42]. This highlights a role for the Tovertafel in enhancing care: by drawing on elements of the games in a social setting, the staff were able to prompt reminiscence by encouraging residents to

share their life stories. Previous research has shown positive outcomes when care staff are involved in reminiscence because they can develop closer relationships with the residents and enhance their care work [46]. We saw this in our study when the staff used the Tovertafel to engage in "biography work," enabling them to learn new things about the residents [33]. Overall, these findings highlight how the Tovertafel can gently prompt reminiscence in ways that benefit both the residents and care staff. They also point to a positive feature of the Tovertafel games, whereby the open-ended design of most games allows for creative facilitation by the staff [34].

However, it is important to note that these benefits were contingent on the significant involvement of the care staff to enable the participation of the residents. On the basis of our findings, it is unlikely that similar outcomes will be achieved if the residents are not supported to use the Tovertafel. This may demand additional staff time in cases where the staff are not already rostered to support enrichment activities. The need for enablement by the staff, coupled with the high cost of the technology itself, should be a key consideration when deciding whether to invest in the Tovertafel. This study suggests that the Tovertafel can be an enriching small-group activity that can meet the needs of individual residents while affording opportunities for creative care work. However, as the staff had discovered at the home we studied, the Tovertafel should be adopted as part of a diverse activities program that includes large- and small-group activities; those that use digital technology; and those that encourage investment in physical infrastructure, such as gardening and interest-based hobbies [10,13].

Organizational Considerations for the Use of the Tovertafel

Our findings highlight several considerations that may inform decisions to implement the Tovertafel within residential aged care.

First, the Tovertafel is a high-cost investment that requires a significant amount of ongoing support to maintain use of the resource. The level of involvement from the staff required to effectively facilitate the Tovertafel is a potential barrier, especially in residential care homes where staff shortage is high, and priority is often given to meeting the personal care needs of the residents [2]. Aligning with previous work [21], our findings emphasize the competencies required by the staff to encourage the residents to use the Tovertafel, especially when those residents have cognitive impairment. This requires a team effort as the Tovertafel appears to work best with support from at least 1 staff member who can facilitate the engagement of the residents and maintain their attention.

Although training staff about the benefits and features of the Tovertafel is important [47], our findings demonstrate that training the staff *how* to creatively facilitate sessions may be even more valuable for maximizing its impact. This type of training is unlikely to be attained via reading manuals or resources and is better acquired through hands-on experience—as was the case with our participants. Training staff on good practices related to the Tovertafel could help with "onboarding" new staff to deliver the activity, given that staff turnover is a known barrier to the successful implementation of technology [16]. Alternatively, training could focus on developing the key competencies of the staff required for facilitation (particularly communication skills), allowing them to facilitate a range of activities [48].

In this study, we observed that facilitation skills were modeled to new staff by experienced staff and mentors. For instance, training typically involved a staff member demonstrating effective strategies while explaining the reasons for each approach aloud. However, this approach is likely to be an obstacle in residential aged care because of high staff turnover and limited staff-family training. Therefore, we recommend that staff receive general communication partner training, which can equip people with the skills required to respond effectively to the needs of the residents and to facilitate communication strategies [49]. Such training focuses on building interpersonal skills and fostering effective interactions, enabling staff to better engage with residents and enhance their facilitation capabilities.

Our study highlighted the need for an appropriate space to enable use of the Tovertafel. In the home we studied, the Tovertafel was situated in 2 specialized care areas: one for people with physical care needs and another for residents with severe cognitive impairment. However, in both locations, the Tovertafel was used in the main dining spaces, where the residents could not always see the projected images clearly. The need to set up the Tovertafel and pack it away meant it was not readily available for spontaneous activity or facilitation by family members and visitors. The residents and the staff felt that the system might be used more often if it were set up in a dedicated space, such as an activities room. This would enable the residents to access the resource with the staff as well as other available facilitators [21].

Third, appropriate scheduling of the Tovertafel needs to be considered to prevent disinterest among residents. According to the staff we interviewed, the Tovertafel had been dormant for some time before this study. This was because the residents had become uninterested in using it, leading to low attendance at the sessions. At the time of this study, staff were reintroducing the Tovertafel as a new activity to elicit interest. However, we observed that the Tovertafel was not well known among the residents compared to other activities. The Tovertafel was challenging to describe and required the staff to promote the benefits of the activity, relying on the residents' prior knowledge and interest in technology. Thus, the staff found it difficult to encourage residents to participate. This is challenging because the Tovertafel is an expensive technology and may be best deployed with larger numbers of residents so that it provides a return on investment.

Studies have reported that facilitators for promoting participation in technology-based activities can include communicating the benefits of use, building knowledge and confidence in using technology, and offering support during its use. In line with these findings, we offer some practical strategies for promoting interest in the Tovertafel: (1) communicating benefits; organizing informative sessions, where staff demonstrate the technology and its benefits, may increase residents' familiarity and engagement; (2) building knowledge and confidence in

using technology; sharing success stories and positive outcomes from previous sessions can serve as motivation for residents to participate; and (3) offering support; a buddy system could be established, pairing residents who are comfortable with technology with those who may be hesitant, fostering a supportive environment for participation. These strategies should be co-designed with residents and staff to ensure they are suitable for this context.

Limitations and Future Work

One limitation of this study is that we were only able to capture data about a handful of Tovertafel sessions, which occurred during a period of 4 weeks with a small number of residents. This may have constrained our observations on the breadth of strategies used by the staff. This limitation was mitigated through the use of diaries that were completed by the staff in additional sessions. Future studies can explore the impact of varying session lengths on the engagement of residents and strategies used by the staff to gain additional insights into effective facilitation techniques. Future research could also monitor the use of Tovertafel over the longer term to ascertain whether sustained engagement by both the staff and the residents occurs. Given the high cost of the device, ongoing use would be required for it to be cost-effective; otherwise, financial resources might be better spent on other beneficial activities and programs with higher engagement over the longer term.

Second, we focused on a single care home in Australia. However, gaining access to residential aged care for research is challenging, and studies are often based on a single site [8,10]. In addition, the Tovertafel is still not in widespread use, making it challenging to identify appropriate study sites to compare findings. Our findings have value in expanding the knowledge base about the Tovertafel and contribute to an improved understanding of the role that staff play in facilitating technologies that are hard for residents to use independently. For example, the benefits of enjoyment and physical activity from the Tovertafel align with a previous study conducted in the United Kingdom [21]. Nevertheless, further research should be conducted to identify context-dependent benefits and evaluate the transferability of our findings.

This study offers limited insight into the views of the residents. Many of the residents observed in our study did not have the cognitive capacity to consent to interviews. The residents who were interviewed had little to say about the Tovertafel, possibly because they were not fully aware of its potential features or because of their brief interaction with it. Future work should aim to provide a deeper understanding of the experiences of the residents. This information could then be used to inform best practices for care staff.

Conclusions

This study investigated the role of the staff in facilitating the effective use of the Tovertafel in 1 Australian residential aged care home. We found that creative strategies were needed to encourage the residents to use the Tovertafel and to support the participation of the residents in the activity. Beneficial outcomes included experiencing fun and joy, social interaction, and reminiscence. The staff proactively mitigated challenges in engagement, which included barriers to access and the adverse reactions that the residents may have to the experience. Aligning with previous research on the use of technology for enrichment in aged care [11,15,17], this study highlights the pivotal role of care staff in the adoption and use of digital technology, meaning that aged care providers need to account for staff involvement when deciding whether to adopt the Tovertafel.

Acknowledgments

This study was funded by the Australian government through a grant from the Australian Research Council awarded to JW (FT170100420) and by the University of Melbourne's Creativity and Wellbeing Research Institute. The authors would like to thank the participants for their willingness to support this study.

Authors' Contributions

JW conceived this study. JW, RMK, and AM designed the study procedures. RMK and AM collected the data. RMK led the data analysis and drafted the manuscript. RMK and AM revised the manuscript in response to reviewers' comments. RO, JAL, and JW edited each version of the manuscript and provided additional feedback.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Design of the feedback diary completed by the staff participants.
[DOCX File , 30 KB - aging_v8i1e67919_app1.docx]

References

1. Older Australians. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/older-people/older-australians/contents/aged-care [accessed 2025-10-11]

2. Monro C, Mackenzie L, O'Loughlin K, Low L, Du Toit SH. 'I could no longer cope at home': experiences of clients and families in residential aged care within the context of Australia's aged care reforms. Australas J Ageing 2023 Jun 28;42(2):344-354. [doi: 10.1111/ajag.13159] [Medline: 36437807]
3. Scott S, Raynor A, Dare J, Grieve J, Costello L. Improving the transition of older adults into residential aged care: a scoping review. Clin Gerontol 2024 Nov 06;47(5):746-759 [FREE Full text] [doi: 10.1080/07317115.2023.2274042] [Medline: 37929882]
4. Health promotion glossary of terms 2021. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240038349 [accessed 2025-05-29]
5. Siette J, Dodds L, Surian D, Prgomet M, Dunn A, Westbrook J. Social interactions and quality of life of residents in aged care facilities: a multi-methods study. PLoS One 2022 Aug 29;17(8):e0273412 [FREE Full text] [doi: 10.1371/journal.pone.0273412] [Medline: 36037181]
6. Miller E. Beyond bingo: a phenomenographic exploration of leisure in aged care. J Leis Res 2017 Dec 13;48(1):35-49. [doi: 10.18666/jlr-2016-v48-i1-6254]
7. Tak SH, Kedia S, Tongumpun TM, Hong SH. Activity engagement: perspectives from nursing home residents with dementia. Educ Gerontol 2015 Mar 21;41(3):182-192 [FREE Full text] [doi: 10.1080/03601277.2014.937217] [Medline: 25489122]
8. Thach KS, Lederman R, Waycott J. Key considerations for the design of technology for enrichment in residential aged care: an ethnographic study. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023 Presented at: CHI '23; April 23-28, 2023; Hamburg, Germany p. 1-16 URL: https://dl.acm.org/doi/10.1145/3544548.3581176 [doi: 10.1145/3544548.3581176]
9. Tierney L, Beattie E. Enjoyable, engaging and individualised: a concept analysis of meaningful activity for older adults with dementia. Int J Older People Nurs 2020 Jun 05;15(2):e12306. [doi: 10.1111/opn.12306] [Medline: 32022414]
10. Thomas JE, O'Connell B, Gaskin CJ. Residents' perceptions and experiences of social interaction and participation in leisure activities in residential aged care. Contemp Nurse 2013 Oct 17;45(2):244-254. [doi: 10.5172/conu.2013.45.2.244] [Medline: 24299253]
11. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162 [FREE Full text] [doi: 10.2196/31162] [Medline: 34975014]
12. Muñoz D, Pedell S, Sterling L. Evaluating engagement in technology-supported social interaction by people living with dementia in residential care. ACM Trans Comput Hum Interact 2022 Oct 20;29(5):1-31. [doi: 10.1145/3514497]
13. Wilding R, Barbosa Neves B, Waycott J, Miller E, Porter T, Johnston J, et al. Introducing virtual reality to older adults: a qualitative analysis of a co-design innovation with care staff. Arch Gerontol Geriatr 2024 Oct;125:105505 [FREE Full text] [doi: 10.1016/j.archger.2024.105505] [Medline: 38851090]
14. Hodge J, Balaam M, Hastings S, Morrissey K. Exploring the design of tailored virtual reality experiences for people with dementia. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC p. 1-13 URL: https://dl.acm.org/doi/10.1145/3173574.3174088 [doi: 10.1145/3173574.3174088]
15. Waycott J, Kelly RM, Baker S, Neves BB, Thach KS, Lederman R. The role of staff in facilitating immersive virtual reality for enrichment in aged care: an ethic of care perspective. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-17 URL: https://dl.acm.org/doi/10.1145/3491102.3501956 [doi: 10.1145/3491102.3501956]
16. Hegney DG, Rees CS, Osseiran-Moisson R, Breen L, Eley R, Windsor C, et al. Perceptions of nursing workloads and contributing factors, and their impact on implicit care rationing: A Queensland, Australia study. J Nurs Manag 2019 Mar 17;27(2):371-380. [doi: 10.1111/jonm.12693] [Medline: 30221436]
17. Kelly RM, Xing Y, Baker S, Waycott J. "More than just holding the iPad": family members' perspectives on the work of video calling in aged care. Proc ACM Hum Comput Interact 2024 Nov 08;8(CSCW2):1-26. [doi: 10.1145/3687007]
18. Cavenett W, Baker S, Waycott J, Carrasco R, Robertson E, Vetere F, et al. Deploying new technology in residential aged care: staff members' perspectives. In: Proceedings of the 30th Australian Conference on Computer-Human Interaction. 2018 Presented at: OzCHI '18; December 4-7, 2018; Melbourne, Australia p. 200-204 URL: https://dl.acm.org/doi/10.1145/3292147.3292214 [doi: 10.1145/3292147.3292214]
19. Zhao W, Steven B, Jenny W. Challenges of deploying VR in aged care: a two-phase exploration study. In: Proceedings of the 32nd Australian Conference on Human-Computer Interaction. 2020 Presented at: OzCHI '20; December 2-4, 2020; Sydney, Australia p. 87-98 URL: https://dl.acm.org/doi/abs/10.1145/3441000.3441018 [doi: 10.1145/3441000.3441018]
20. The power of purposeful play. Tover. URL: https://www.tover.care/au/ [accessed 2024-10-22]
21. Good A, Omisade O, Ancient C, Andrikopoulou E. The use of interactive tables in promoting wellbeing in specific user groups. In: Proceedings of the 5th International Conference on Human Aspects of IT for the Aged Population, Held as Part of the 21st HCI International Conference. 2019 Presented at: HCII '19; July 26-31, 2019; Orlando, FL p. 506-519 URL: https://link.springer.com/chapter/10.1007/978-3-030-22015-0_39 [doi: 10.1007/978-3-030-22015-0_39]
22. Talman L, Gustafsson C. Evaluation of Tovertafel UP. Academy of Health, Care and Social Welfare. 2020. URL: https://www.diva-portal.org/smash/get/diva2:1415974/FULLTEXT01.pdf [accessed 2025-05-29]
23. Why is research so important? Tover. URL: https://www.tover.care/au/research [accessed 2024-10-22]

24. Bruil L, Adriaansen MJ, Groothuis JW, Bossema ER. [Quality of life of nursing home residents with dementia before, during and after playing with a magic table]. Tijdschr Gerontol Geriatr 2018 Apr 12;49(2):72-80. [doi: 10.1007/s12439-017-0243-3] [Medline: 29235073]
25. Kelly RM, Cheng Y, McKay D, Wadley G, Buchanan G. “It’s about missing much more than the people”: how students use digital technologies to alleviate homesickness. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan p. 1-17 URL: https://dl.acm.org/doi/10.1145/3411764.3445362 [doi: 10.1145/3411764.3445362]
26. Kelly RM, Xing Y, Baker S, Waycott J. Video calls as a replacement for family visits during lockdowns in aged care: interview study with family members. JMIR Aging 2023 Jun 12;6:e40953 [FREE Full text] [doi: 10.2196/40953] [Medline: 37191951]
27. Baker S, Waycott J, Robertson E, Carrasco R, Neves BB, Hampson R, et al. Evaluating the use of interactive virtual reality technology with older adults living in residential aged care. Inf Process Manag 2020 May;57(3):102105. [doi: 10.1016/j.ipm.2019.102105]
28. Meyer C, Ogrin R, Golenko X, Cyarto E, Paine K, Walsh W, et al. A codesigned fit-for-purpose implementation framework for aged care. J Eval Clin Pract 2022 Jun 06;28(3):421-435 [FREE Full text] [doi: 10.1111/jep.13660] [Medline: 35129259]
29. Spilsbury K, Haunch K, Hodge G, Pearce S. The opportunities and challenges of conducting observational research in care homes: what a researcher brings, does, and leaves behind? J Long Term Care 2024;63(2):360-370 [FREE Full text] [doi: 10.31389/jltc.280]
30. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: Sage Publication; 2022.
31. Terry G, Hayfield N. Essentials of Thematic Analysis. Washington, DC: American Psychological Association; 2021.
32. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2020 Aug 12;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
33. Müller C, Neufeldt C, Randall D, Wulf V. ICT-development in residential care settings: sensitizing design to the life circumstances of the residents of a care home. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 2639-2648 URL: https://dl.acm.org/doi/10.1145/2207676.2208655 [doi: 10.1145/2207676.2208655]
34. Waycott J, Davidson J, Baker F, Yuan S. Supporting creativity in aged care: lessons from group singing, music therapy, and immersive virtual reality programs. In: Proceedings of the 34th Australian Conference on Human-Computer Interaction. 2022 Presented at: OzCHI '22; November 29-December 2, 2022; Canberra, Australia p. 272-282 URL: https://dl.acm.org/doi/abs/10.1145/3572921.3572940 [doi: 10.1145/3572921.3572940]
35. Gerling KM, Mandryk RL, Linehan C. Long-term use of motion-based video games in care home settings. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 1573-1582 URL: https://dl.acm.org/doi/10.1145/2702123.2702125 [doi: 10.1145/2702123.2702125]
36. Dahl Y, Löfström E. Supporting social interaction in care environments: exploring stakeholder perspectives on the potential of interactive technology. Int J Hum Comput Interact 2018 Jan 24;35(1):53-64. [doi: 10.1080/10447318.2018.1426897]
37. Barbosa Neves B, Sanders A, Kokanović R. "It's the worst bloody feeling in the world": experiences of loneliness and social isolation among older people living in care homes. J Aging Stud 2019 Jun;49:74-84. [doi: 10.1016/j.jaging.2019.100785] [Medline: 31229221]
38. Blythe M, Wright P, Bowers J, Boucher A, Jarvis N, Reynolds P, et al. Age and experience: ludic engagement in a residential care setting. In: Proceedings of the 8th ACM Conference on Designing Interactive Systems. 2010 Presented at: DIS '10; August 16-20, 2010; Aarhus, Denmark p. 161-170 URL: https://dl.acm.org/doi/10.1145/1858171.1858200 [doi: 10.1145/1858171.1858200]
39. Gaver W, Boucher A, Bowers J, Blythe M, Jarvis N, Cameron DW, et al. The photostroller: supporting diverse care home residents in engaging with the world. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC p. 1757-1766 URL: https://dl.acm.org/doi/10.1145/1978942.1979198 [doi: 10.1145/1978942.1979198]
40. Downs J, Frank V, Wally S. Differentiated participation in social videogaming. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. 2015 Presented at: OzCHI '15; December 7-10, 2015; Parkville, Australia p. 92-100 URL: https://dl.acm.org/doi/10.1145/2838739.2838777 [doi: 10.1145/2838739.2838777]
41. Henkel LA, Kris A, Birney S, Krauss K. The functions and value of reminiscence for older adults in long-term residential care facilities. Memory 2017 Mar 10;25(3):425-435. [doi: 10.1080/09658211.2016.1182554] [Medline: 27161334]
42. Baker S, Kelly R, Waycott J, Carrasco R, Bell R, Joukhadar Z, et al. School's back: scaffolding reminiscence in social virtual reality with older adults. Proc ACM Hum Comput Interact 2021 Jan 05;4(CSCW3):1-25. [doi: 10.1145/3434176]
43. Porock D, Cain P, Young C, Hilbers J, Bomm M, Ghosh M, et al. The careful project: evaluating the acceptability of local virtual reality experiences as a leisure activity for residents in aged care. Australas J Ageing 2025 Mar 07;44(1):e13379. [doi: 10.1111/ajag.13379] [Medline: 39370985]
44. Webber S, Baker S, Waycott J. Virtual visits: reminiscence in residential aged care with digital mapping technologies. Australas J Ageing 2021 Sep 04;40(3):293-300. [doi: 10.1111/ajag.12902] [Medline: 33538405]

45. Cooney A, O'Shea E, Casey D, Murphy K, Dempsey L, Smyth S, et al. Developing a structured education reminiscence-based programme for staff in long-stay care facilities in Ireland. J Clin Nurs 2013 Jul 14;22(13-14):1977-1987. [doi: 10.1111/j.1365-2702.2012.04342.x] [Medline: 23241022]
46. Perion JJ. Tovertafel: evaluating the benefits of a novel multi-sensory intervention for nursing home residents with dementia [thesis]. The University of Toledo. 2021. URL: http://rave.ohiolink.edu/etdc/view?acc_num=toledo1620142242101885 [accessed 2025-08-04]
47. Bennett M, von Treuer K, McCabe MP, Beattie E, Karantzas G, Mellor D, et al. Resident perceptions of opportunity for communication and contribution to care planning in residential aged care. Int J Older People Nurs 2020 Mar 02;15(1):e12276. [doi: 10.1111/opn.12276] [Medline: 31578823]
48. Burton B, Shrubsole K, Manchha A, King M, Wallace SJ. Communication partner training for aged-care workers: a scoping review. Int J Lang Commun Disord 2025 Feb 20;60(2):e70016. [doi: 10.1111/1460-6984.70016] [Medline: 39977832]
49. Benoit-Dubé L, Jean EK, Aguilar MA, Zuniga A, Bier N, Couture M, et al. What facilitates the acceptance of technology to promote social participation in later life? A systematic review. Disabil Rehabil Assist Technol 2023 Apr 06;18(3):274-284. [doi: 10.1080/17483107.2020.1844320] [Medline: 33156714]

Abbreviations

PICF: participant information sheet and consent form
VR: virtual reality

Edited by J Wang; submitted 25.10.24; peer-reviewed by V Provencher, D Healy; comments to author 14.01.25; revised version received 04.06.25; accepted 03.07.25; published 12.08.25.

Please cite as:
Kelly RM, Manchha A, Waycott J, Ogrin R, Lowthian JA
Staff Enablement of the Tovertafel for Enrichment in Residential Aged Care: Field Study
JMIR Aging 2025;8:e67919
URL: https://aging.jmir.org/2025/1/e67919
doi:10.2196/67919
PMID:40794418

Exploring Technology Supporting Aging-in-Place Using an Equity Lens Through Focus Groups and World Café–Informed Research Agenda: Qualitative Study

Marianne Saragosa[1,2*], PhD; Ibukun-Oluwa Omolade Abejirinde[3,4,5*], MSc, MD, PhD; Evan MacEachern[6], BKin, MScRR; Michelle LA Nelson[2,6*], PhD; Kristina M Kokorelias[7,8,9], PhD; Sidra Bharmal[1*], BSc; Brina Ludwig Prout[10*], BA; Marian Mohamed[1*], BSCN

[1]Science of Care Institute, Sinai Health, 1 Bridgepoint Drive, Toronto, ON, Canada

[10]Hennick Bridgepoint Hospital, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[3]Institute for Better Health, Trillium Health Centre, Mississauga, ON, Canada

[4]Research and Innovation Institute, Women's College Hospital, Toronto, ON, Canada

[5]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[6]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada

[7]Department of Geriatric Medicine, Sinai Health System and University Health Network, Toronto, ON, Canada

[8]Department of Occupational Science and Occupational Therapy, Temerty Faculty of Medicine, Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada

[9]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Marianne Saragosa, PhD
Science of Care Institute, Sinai Health, 1 Bridgepoint Drive, Toronto, ON, Canada

Abstract

Background: Older adults prefer to age in their home or community of choice, which could include naturally occurring retirement communities (NORCs). As a place with a high density of older adults, NORCs could be sites where technology is leveraged to support independence and aging in the right place. However, there is limited research on how technology adoption and use occur in NORCs in ways that support older adults.

Objective: This study aims to cocreate a research agenda on equity-informed technology considerations that help older adults live independently in NORCs.

Methods: This is a 2-phase sequential qualitative descriptive study of 5 community-based focus groups and an in-person World Café event. We use the focus group method to acquire data about older adults' experiences with and perceptions of using technology to support aging-in-place in NORC settings. This data informs the design and facilitation of deliberate dialogues at the World Café event. Three questions helped to guide the small group discussions. The World Café is a creative, collaborative, and conversation-generating method that aims to generate exchanges between people with different views on a particular topic.

Results: In total, 45 NORC residents participated in a focus group about their experience and use of technology. The data revealed 3 central categories that highlight the perception of the use of technology to support the independence of participants in their homes and communities, its challenges, and areas to consider when deploying technology for helping older adults age in place. The subsequent World Café event included 40 participants and a combination of NORC residents, service providers, researchers, technology innovators, and policy makers. Insights drawn from the focus groups and World Café informed a 10-question research agenda about equity-informed technology principles that span accessible support, accessible interfaces, affordable and equitable access, available digital literacy training, accessible data, and accessible partnerships.

Conclusions: Our study explores NORCs as potential environments for offering a transformative opportunity to address equity considerations for technology supporting aging in place. Our findings and research agenda highlight critical areas for consideration, including leveraging partnerships, integrating public and private technology ecosystems, and designing technology with older users that evolves with the population's needs. Notably, by embedding principles of equity, inclusivity, and user-centered design, the collective of developers, researchers, and service providers can ensure that emerging technology serves diverse aging populations equitably and effectively.

(JMIR Aging 2025;8:e71093) doi:10.2196/71093

KEYWORDS

aging in place; community; research agenda; older adult; world cafe; technology; equity

Introduction

Background

As the fastest-growing age demographic in Canada, older adults are expressing a strong desire to age in place [1]. This research defines an "older adult" as a person aged 65 years or older [2]. Aging in place refers to living safely, independently, and comfortably in one's home and community regardless of age, income, or ability level [3]. However, "aging in the right place" has been considered a more appropriate goal, given that place often depends on an older person's personal preference, circumstances, and care needs [4]. Advocates suggest expanding housing options for those unsuitable for long-term care by developing alternatives, such as senior apartments, assisted living residences, and so forth to create safer and more accessible living environments and support the desires of older adults to age in place [4,5]. Such settings could include adaptations made to the residential environments of older people for supporting their independence and improving their quality of life [6]. For example, incorporating digital technologies, such as movement sensors and medication dispensers, is increasingly used to mitigate declining cognitive and physical abilities [7].

One example of an alternative housing model is naturally occurring retirement communities (NORCs). Originally coined in 1986, NORCs are defined as a "housing development that is not planned or designed for older people, but which over time comes to house largely older people" [8]. The NORC Innovation Centre at University Health Network has expanded the conceptualization of NORCs to include that they may consist of communities designed to house many older adults but not purpose-built for this population in the way long-term care homes or retirement homes are [9]. NORCs can provide a model for healthy aging-in-place since many offer supportive service programs based on the needs of residents [8]. While the availability and capacity of government, social, and health service providers and policy makers can create opportunities for supportive programming [10], less is known about how technology is currently being leveraged within NORC environments to support the needs of residents. Technological advancements (eg, remote monitoring systems, wearable health devices, and mobile apps) have shown promise in supporting efforts to age in place [11,12]. For example, countries like the Netherlands and Spain have focused on technology to mitigate the need for additional care and promote independence and autonomy [13]. In particular, low-tech technologies are designed to be as simple as possible and can offer practical, low-cost, and accessible features [14]. Despite considerable public and private interest and investment in aging-in-place technology development and research [15,16], there has been limited adoption of advanced technologies and their impact on people and health care systems [17], including diverse older adults [18,19]. Gaps persist in accessing technology and enabling digital literacy to support aging in the place where one prefers [20].

Later in life, technology acceptance and adoption are often influenced by the digital divide, which is known as the gap between those who do and those who do not have access to or knowledge of how to use advanced forms of technology [21]. While many older adults actively use current technologies (eg, internet, tablets, and smartphones), the digital divide continues to be a barrier for older people with lower access (ie, devices, internet, financial, and digital literacy) [22]. The digital divide is even more pronounced in those who experience marginalization at the intersections of racial and ethnic identity and socioeconomic status [23-25]. Indeed, significant barriers to accessing or engaging with technologies for health-related needs could worsen racial and ethnic health disparities experienced by the aging population [25]. While technology can offer promising solutions to support older adults in Canada and other countries in aging in the right place [19], the digital divide, or the gap between those who do and those who do not have access to new digital and information technology, must be addressed [21].

Objective

With limited research on how technology adoption and use occur in NORCs in ways that support aging in place, this study aimed to cocreate a research agenda on equity-informed technology considerations that help older adults live independently in NORCs.

Methods

Study Design

This 2-phase sequential qualitative descriptive study [26] consisted of community-based in-person focus groups and an intersectoral World Café event. We used the focus group method to gather data about older adults' experiences with and perceptions of using technology to support aging-in-place in NORC settings. The focus group data, in combination with project team discussions, helped develop the overarching questions for the deliberate dialogues at the World Café event [27]. The World Café is a creative, collaborative, and conversation-generating method that aims to generate exchanges between people with different views on a particular topic [28].

Developed in 1995, the World Café is considered a participatory method [29] by shifting to a bottom-up approach by engaging as many different actors as possible in the data collection process [30]. The World Café was the preferred approach in this case, as it has been used extensively to convene community-dwelling older adults and intersectoral participants [31,32], including research prioritization activities [28]. The evidence also supports the World Café as a valuable tool for fostering productive conversations among individuals with lived experience. Finally, with World Cafés, a large and heterogeneous group of people

can be brought together in a systematic and organized manner [30].

We used the following 5 steps of the World Café method to guide our event planning: creating an informal environment, offering a warm welcome and overview from the host, engaging in 3 rounds of conversations among smaller groups of participants on 3 key questions (see), and a final round of sharing the conversation results with the broader group [29]. Typically, participants are encouraged to rotate tables intermittently and meet new people; however, since some participants had mobility differences, we adapted the format such that the "tabletop host" and a notetaker (ie, scribe) rotated tables instead. The following were the World Café questions:

1. What challenges do older adults living in naturally occurring retirement community (NORC) neighborhoods have that might be supported with technology to help them stay in their homes?
2. How might challenges with technology be the same or different for people who are newcomers, are from different cultural backgrounds, education, and income levels?
3. What other things should we think about when developing and using technological (or digital) tools that could help older people stay in their homes?

Sample and Recruitment

Focus Groups

We used purposeful sampling to identify and invite older adults living in urban NORCs in Toronto, Canada, to participate in a focus group. The inclusion criteria were community-dwelling older adults living in a NORC building. We identified NORCs from a publicly available list, including identified apartments, condos, co-operative housing, and social housing buildings within a bounded geographical area [33]. We used personal and professional networks to approach NORC buildings about the study. With agreement and support from the condo board or service providers at each location, we invited residents through word-of-mouth, circulating a recruitment flyer and using gatekeepers (ie, NORC residents or managers who brokered relationships between the researchers and the NORC building).

World Café

We also used purposeful sampling to invite interested focus group participants to the World Café. To ensure adequate representation of NORC residents in numbers, we used snowball sampling, encouraging focus group participants to share information about the World Café event with other residents. Because of the World Café's intersectoral nature, we invited representatives from housing, health and social care, academics and research, and government senior services to attend.

Data Collection

Focus Groups

Focus groups were held on-site in communal activity rooms within 5 different NORC buildings between April 2024 and June 2024. Each focus group involved a facilitator and a cofacilitator who supported the consent process, collected sociodemographic information, took field notes, and provided follow-up questions. The facilitator started each session by welcoming the participants, explaining the discussion's purpose, and inviting them to introduce themselves. With an interview guide, the facilitator guided the discussion to address the following three central questions: (1) tell me about your current care needs; (2) tell me how technology addresses or could address these needs; and (3) tell me about current challenges and benefits with technology. For the focus group participants, we broadly defined "technology" as the various tools, devices, and systems people use daily to facilitate tasks, improve efficiency, and enhance convenience, spanning assistive devices, telemedicine, health monitoring, information and communication technologies, and so forth [34]. The focus groups lasted approximately 45-70 minutes; the discussions were digitally audio-recorded and transcribed verbatim by one of the authors (MM).

World Café

The in-person, full-day World Café event occurred in June 2024 in a large workspace at a tertiary hospital in Toronto, Canada. Upon arrival, participants were asked to complete paperwork (eg, consent forms and demographic survey) and were seated at one of six tables. Each table had a tabletop host and a notetaker. The host and notetaker were assigned one of three questions and rotated at 3 intervals to allow all participants to build on the prior responses to the question and provide additional feedback. Upon completion of the third round, all participants were invited to participate in a group discussion, answering the final question: "What groups need to work together to ensure that new technologies that help older people stay in their homes treat everyone fairly?"

Data Analysis

Focus Groups

We used conventional content analysis to identify cross-cutting categories across the focus groups [35]. Conventional content analysis is most beneficial in the absence of existing research on the phenomenon. Therefore, it uses inductive logic, avoiding applying any predetermined categorical interpretation system [35]. Data analysis started with reading all transcripts, and 2 coders identified salient codes by highlighting exact words from the text to capture thoughts or concepts that appeared to describe an experience or perception of technology within the context of aging in place. Next, the first author created labels that reflected more than one concept, thereby sorting and merging codes into categories. Finally, the emergent categories were used to organize and group all the codes into meaningful clusters. For example, quotes like "[using technology to] connect with people near and far" and "interacted with all the kids and grandkids" were coded as "communication" and categorized under "Uses of technology."

World Café

For the World Café event, we used directed content analysis. This analysis seeks to validate or extend a theory or framework [35]. In this case, we applied the Social Ecological Model (SEM) [36] as an organizing framework to map the factors of technology use among NORC residents, considering the challenges and enablers. The SEM was selected because it

suggests that behavior is integrated into a dynamic network of intrapersonal characteristics, interpersonal processes, and institutional, community, and public policy factors [37]. Using the SEM also allows for identifying specific drivers of technology use or disuse among community members. Handwritten notes were transcribed and consolidated into a Microsoft Word document. In the coding process, we included examples for each of the factors. For instance, older adults' knowledge, attitudes, and behaviors are at the "individual" level of the SEM model. Similar to the focus group analysis, we reviewed the transcribed notes and coded the data using predetermined categories wherever possible. For example, better integrating "sensors in NORC buildings" and coordinating specialized technology training, as a NORC was coded according to the SEM's community level, acknowledges these factors as potential facilitators of aging-in-place. We also coded data that could not be categorized into one of the SEM factors, such as data related to technology devices and their features, design, and interface.

Ethical Considerations

This study was approved by the Mount Sinai Hospital Research Ethics Board (REB reference number 23‐0123-E). Study participants provided written informed consent before engaging in study activities. Three participants in each focus group received a $25 CAD (US $18) gift card as a draw, whereas World Café attendees were offered a $25 CAD (US $18) gift card with the option to decline. Participation was not anonymous due to the focus group and in-person event; however, participation was voluntary. The study findings will be disseminated through presentations at conferences and peer-reviewed publications using deidentified data.

Results

Characteristics of Participants

Focus Groups

In total, 45 people participated in a focus group, with each focus group having between 6 and 25 NORC residents. Of the demographic data we collected, most of the participants identified as female (n=30, 67%), between 70 and 89 years of age (n=33, 74%). The most identified ethnicities were White North American or European (n=29, 66%), Southeast Asian (n=3, 7%), and East Asian (n=3, 7%). Most participants indicated living alone (n=27, 60%), followed by living with a partner (n=14, 31%). The remaining participants lived with others (including family) or preferred to keep their information private. See Table 1 for Focus group participant characteristics.

Table . Focus group participant characteristics.

Characteristic	NORC[a] resident (n=45)
Gender, n (%)	
Women	30 (67)
Men	15 (33)
Age (y), n (%)	
60 - 69	10 (22)
70 - 79	21 (47)
80 - 89	12 (27)
90 - 99	1 (2)
Prefer not to answer	1 (2)
Ethnicity, n (%)	
White Northern American	19 (42)
White European	10 (22)
East Asian	3 (7)
Southeast Asian	3 (7)
Latino or Hispanic	2 (4)
Caribbean	1 (2)
Black African	1 (2)
Middle Eastern	1 (2)
Mixed heritage	1 (2)
Indigenous	1 (2)
Prefer not to answer	3 (7)
Living situation, n (%)	
Live alone	27 (60)
Live with partner	14 (31)
Live with family or friends	2 (4)
Prefer not to answer	2 (4)
Relationship status, n (%)	
Single	22 (49)
In a relationship	13 (29)
Widow	3 (7)
Separated	2 (4)
Prefer not to answer	5 (11)
Employment status, n (%)	
Retired	37 (82)
Unemployed, looking for work	3 (7)
Employed for wages	1 (2)
Prefer not to answer	4 (9)

[a]NORC: naturally occurring retirement community.

World Café

A total of 40 people participated in the World Cafe. Of these, 18 identified themselves as NORC residents; the remaining individuals were system leaders, including service providers, decision-makers, researchers, technology entrepreneurs, innovators, and administrators. Of the NORC residents, most identified as female (n=14, 78%), were between the ages of 70

and 79 (n=9, 50%) years, lived alone (n=9, 50%) or with a partner or family (n=8, 44%), and most identified as White European or North American (n=12, 66%). For the system leaders (N=22), the average years of experience was 11 (SD 4.2) years, ranging from 2 to 17 years. Most participants held a graduate degree (n=15, 68%) and identified as female (n=18, 82%). See Table 2 for World Café participant characteristics.

Table . World Café participant characteristics.

Characteristics	Role	
	NORC[a] resident (n=18)	System leader (n=22)
Gender, n (%)		
Female	14 (78)	18 (82)
Male	4 (22)	4 (18)
Age (y), n (%)		
18 - 29	0 (0)	1 (4)
30 - 39	0 (0)	7 (32)
40 - 49	0 (0)	11 (50)
50 - 59	0 (0)	1 (4)
60 - 69	4 (22)	0 (0)
70 - 79	9 (50)	1 (4)
80 - 89	4 (22)	0 (0)
90 - 99	1 (6)	0 (0)
Prefer not to answer	0 (0)	1 (4)
Ethnicity, n (%)		
White European	8	5 (23)
White North American	4 (22)	5 (23)
East Asian	1 (6)	3 (14)
Southeast Asian	1 (6)	1 (4)
Caribbean	2 (11)	1 (4)
Black African	0 (0)	1 (4)
Middle Eastern	0 (0)	1 (4)
Mixed heritage	0 (0)	2 (9)
Prefer not to answer	2 (11)	3 (14)
Living situation, n (%)		
Live alone	9 (50)	—[b]
Live with partner	8	—
Live with family	1 (6)	—
Relationship status, n (%)		
In a relationship	8	—
Single	6 (33)	—
Widow	1 (6)	—
Prefer not to answer	3 (17)	—
Employment status, n (%)		
Retired	15 (83)	—
Unable to work	1 (6)	—
Self-employed	1 (6)	—
Prefer not to answer	1 (6)	—
Years of experience		
Mean (SD)	—	10.8 (4.2)
Range	—	2 - 17

Characteristics	Role	
	NORC[a] resident (n=18)	System leader (n=22)
Profession, n (%)		
Researcher		11(50)
Service provider	—	4 (18)
Designer/innovator	—	4 (18)
Administrator	—	2 (9)
Policy maker	—	1 (5)
Education level completed, n (%)		
Postsecondary	—	6 (27)
Diploma	—	1 (5)
Graduate degree	—	15 (68)

[a]NORC: naturally occurring retirement community.

[b]Not applicable.

Focus Group Findings

The focus group data revealed 3 central categories that highlight the use of technology to support participants' independence in their homes and communities, its challenges, and areas to consider when deploying technology to help older adults age in place (see Table 3).

Table . Focus group findings.

Category	Open codes	Direct quotes
Using technology	• Life is online • Connecting everyday • Finding information	• "When I leave my appointment everything comes online…medical appointments, email, the whole world." [FG 3] • "If there's something I don't understand how to do, I Google it and it sends me to YouTube where it explains how to do it. But it's so fast-paced that I have to keep rewinding it." [FG 2]
Expressing challenges using technology	• Goofing up the machine (cell phone) • Remembering passwords	• "I do not have a cell phone. I got rid of it a year and a half ago… I can goof the machine up very easy. I always seem to do that with cell phones." [FG 2] • "You have to remember to change your password every six months. We locked out and out because we forgot it." [FG 4]
Identifying technology considerations	• Needing to be taught • Thinking about income	• "My emails right now are coming into trash. And I'm at the point where I won't touch anything." [FG 2] • "We have been waiting for the whole building to be so-called fiber. Well, unfortunately, they have been trying to entice me to join by asking me to spend $90.00 a month. For a retiree, it's not really that simple" [FG 4]

Using Technology

Focus group participants described how they currently use technology to support their independence in their homes and communities. The types of use fell into categories of communication (eg, keeping in contact with family and friends), wellness and activity (eg, tracking upcoming appointments and online banking), information seeking (eg, accessing news, information, and current events), and for personal safety (eg, checking in on others). Some participants recognized the role of existing (eg, printer and video calling) and emerging technology (eg, ChatGPT and wearables) in adding value to their daily activities through access to information and connection to others.

> *I have a new iPhone. And I wanted to use it to make films. So I figured to make a film, you have to have a*

story to start. So, I asked ChatGPT what story I could use. And it gave me a story that was much, much too complicated. So you have a conversation [with ChatGPT] about simplifying it and making it something you can do in two-minute films? [FG 1]

Expressing Challenges Using Technology

Cross-cutting challenges expressed by the participants included the cost associated with technology, such as data or internet plans or fall alert technology; physical changes that make using technology more complex; and concerns related to fraud and data privacy. The latter was expressed in all focus groups, highlighting the vulnerability felt by older adults to financial scams. Another challenge often mentioned was the need for more digital skills or the inability to use technology, such as cell phones, optimally. This challenge was well aligned with the expectation that older adults must become familiar with current technology, as many tasks related to day-to-day societal functioning have become digitalized. This includes accessing lab results or medical records and online banking. Several participants also disclosed individual physical and sensory changes that made navigating technology interfaces more difficult for them:

I think some of us have physical problems. I have a lot of vision problems, so sometimes I can't even do anything at all because I can't see where the arrow is… the other thing is that I'm losing the ability to grasp with my hands, and that is horrible, but it, you know, it makes it difficult [FG 4]

Identifying Technology Considerations

Considering the challenges reported by our sample of NORC residents, many participants often suggested how to enhance their experience with technology. For example, they were intentional about receiving more technology support through education, training, and workshop-style learning. Our sample strongly desired someone to walk them through technology (eg, a digital navigator), especially when devices are seemingly not working and require troubleshooting advice. The desire for education also extended to avoiding cyber scams and using emerging technology, like ChatGPT. However, some participants resisted entirely internet-based systems and wondered about having either nontech or low-tech alternatives or technology that considered the unique needs of older adults, for example, a slower typing speed. Finally, participants also mentioned low-cost technology options aligned with their budgetary realities.

World Café Findings

According to the workshop participants, the SEM framework helped highlight challenges and facilitators of technology at the individual, interprofessional, community, organizational, policy, and technology device levels (see Table 4).

Table . Social Ecological Model framework.

Level	Challenge	Facilitator
Individual (ie, knowledge, attitudes, and behaviors)	• Fatigue with online appointments and expectation of online solutions (ie, "Zoom fatigue") • Hesitancy in asking for help with technology • Need for repeated exploration to acquire and develop digital literacy ("One-and-done doesn't work") • One's formal education does not equate to digital literacy	• Desire to have a "real" interaction • Identified characteristics of those who adapt to technology: healthier, open-minded, able to afford technology-related costs, willingness to learn and acknowledge technology shortcomings
Interpersonal (ie, family, friends, and social networks)	• Customer service supporting access to technology is not always respectful and informative • Technology cannot replace human connection	• Family available who can help facilitate set-up; family available to attend virtual meetings • Free of charge or nominal cost-required services equipped to help older people stay connected (eg, Geek Squad and Toronto Public Library)
Community (ie, relationships between organizations)	• Concerns about fraud and scams targeting older adults • Virtual programming may result in challenges (ie, long waitlists, lack of relief to caregivers, or social participation for users) • Ageism toward older people and technology, including them learning how to use technology	• Technology to see what resources or assets are available in the community (eg, grocery delivery and volunteer services) • Power in organizing a "collective idea or request" that supports collective learning and support for integrating technology solutions into living environments • Having people in your building or community who can help and whom you can trust • Building community and technology support ("economies of scale"), which may foster social engagement and participation • Leverage existing, trusted institutions like public libraries
Organizational (ie, organizations and social institutions)	• Customer service or technology support lacks the skills to communicate effectively with older adults (eg, they speak too loudly and give unclear instructions); people would rather live with the problem than deal with technology support	• Technology tutorials or programs (eg, how to sign a document) • Assistive devices in the home (eg, light sensors, passive sensors, and cane sensors) • Existing support system (eg, NORC[a] community coordinator) • Creation of a "digital navigator" to facilitate solution finding and increase user participation
Policy (ie, national, provincial, and local laws)	• Lack of policy requiring technology designed and customized to the needs of older users who consider sensory or cognitive differences • The government seemingly not engaged in supporting use of technology among older adults	• Affordable NORC phone, cell phone, and internet packages • Having technology partners represented at the policy level
Technology device (ie, features, interface, and design)	• Design does not meet the user needs, who may include older adults • Language translation, colors, and fonts need to be unavailable or easily customizable for everyone (eg, bigger font and different colors)	• Use of plain language or avoidance of acronyms • Considering aging-related changes in technology design (eg, mouse speed and color contrast for vision)

[a]NORC: naturally occurring retirement community.

The workshop discussions revealed that technology is not generally tailored to meet the needs of older adults, and the aging-oriented devices that do exist can be improved through more thoughtful and unique designs. According to the workshop participants, co-design is needed, where technology developers work with older adults from the onset of the conceptualization

and design process. Technology should not only be adapted for older adults but also specifically for different demographics and complex needs (eg, hearing impairment). An example of an improvement would be if device applications included multilingual options. Diverse groups of older adults must be consulted in technology development. Therefore, accessibility was frequently discussed, stressing the need for equity in technology design and access. In addition, the lack of affordability of technologies posed a significant barrier to accessibility, significantly disadvantaging lower-income groups. Participants suggested possible improvements, including reduced internet plans or group packages for NORCs.

The widespread adoption of technology to replace traditional services is becoming the default and can complicate many processes for older adults (eg, banking). Instead, there should continue to be a nononline option in addition to technology. The widespread use of technology is often blamed for reducing meaningful human interaction, which is concerning because people naturally seek social connection. Conversely, other participants voiced that information and communication technology allows connecting with others more easily, like family abroad, especially during scenarios of restricted movement, as in the case of the COVID-19 pandemic. Communication technologies (ie, Zoom [Zoom Communications] and Facetime [Apple Inc]) benefit virtual care, where technology can increase access to health care professionals and reduce wait times. Many participants noted they required more support for digital literacy to be comfortable navigating certain technologies, presenting a need for a tutor or someone available to answer technology-related questions, echoing insights from the focus groups.

It was suggested that aging communities, such as NORCs, should partner with programs or community agencies to support older adults in being technologically proficient. Technological literacy is critical in providing NORC residents with user confidence and helps to reduce anxiety around possible scams or security breaches. It was also suggested that instead of being compelled to accept the terms and conditions, users should be able to dictate the information that technology platforms share and with whom. Concurrently, technology could also provide a safety net at home. Many participants already appreciated using technology in remote monitoring and risk identification. Therefore, the ultimate goal to keep in mind when developing technology is how it can provide essential support in maintaining the independence of older adults to age in their preferred environment, which also evolves over an individual's lifetime.

Synthesizing insights drawn from the focus groups and World Café informed a research agenda (Table 5) of questions mapped to equity-informed technology principles.

Table . Research agenda.

Principles	Research question
Accessible support	• How can the NORC[a] environment support access and use of technology among older residents by leveraging existing and new partnerships (ie, City of Toronto, Toronto Public Library, and Connected Canadians)? • How do we leverage public (eg, Remote Care Monitoring) and private technology (eg, Apple Watch) ecosystems to support NORC residents to age in place? • What might be the most impactful ways in which technology could support community building and participation among NORC residents, local service providers, and agencies?
Accessible interfaces	• What are the technology design principles that consider the evolving needs of older adults, and how can these be integrated into current and future applications?
Affordable and equitable access	• How can technology developers be guided to commit to equity and adhere to it during the design and deployment phases of emerging technology? • How can government enact equitable and inclusive policies and funding to support the adoption of senior-friendly technology (ie, AgeTech)? Or policies to support alternatives to technology?
Available digital literacy training	• How do we create community-based co-located senior-friendly technology education and training opportunities to meet the diverse needs of aging adults? • How can education and training offerings be targeted to certain populations, for example, older adults who are new to Canada, those who are multilingual, are underhoused, and persons who have complex care needs?
Accessible data	• How do we measure the impact of technology on NORC residents, and find data-driven solutions?
Accessible partnerships	• How do we create an intersectoral and collaborative network of partners to support NORC residents to access and use technology, and who should be part of the network?

[a]NORC: naturally occurring retirement community.

Discussions

Principal Findings

We conducted a 2-phase sequential qualitative descriptive study of community-based focus groups and a World Café event to develop a research agenda on older adults' informed considerations for technology to support the aging-in-place of persons living in NORCs. We found several cross-cutting categories across study activities. Participants described using technology for various reasons, including connecting with others, wellness and activity, information seeking, and personal safety. However, our sample's technology usage was also impacted by barriers, such as limited access to (or outdated) devices, the cost associated with technology and its use (eg, internet), and a lack of available training or support to help fully use available technology or address low digital literacy within the context of age-related changes (ie, sensory, cognitive, and functional) that could hinder ease of adoption. The final set of 10 potential research questions underscores the importance of creating opportunities for age-friendly technology education, training, and equity in design, including design principles that reflect the evolving needs of older adults. Finally, the questions also emphasize the need to leverage intersectoral partnerships for supporting NORC residents who inhabit a unique living situation that enables the delivery of support within an existing community of persons aging in place. Therefore, accessing and using technology can serve as a mechanism for community building among NORCs and service providers.

Comparison With Prior Work

Our findings on the barriers to using technology among older NORC residents are similar to other research that found challenges across intrinsic (including individual physical and sensory changes) and extrinsic (including inexperience with technology and cost to access) factors [38]. However, some sociodemographic factors may significantly impact specific groups of older adults, like those on lower or fixed incomes, females, ethnic minority groups, those with lower education levels, and those who are older [39,40]. While our study did not examine the intersectional impact of race and sex on access to technology, other research suggests a significant disparity, with non-White older females having the least access to technology [40]. These findings speak to the ongoing digital

divide, referring to the gap between those with access to technology and those without. According to the literature, significant dimensions of the digital divide include digital literacy, affordability (costs), equity-denied group-related content and services (culturally appropriate content, activities, programs, and services), and access (infrastructure) [41].

NORCs, which include buildings or neighborhoods with a high density of older people, are home to highly diverse and socially and medically complex aging demographics [42,43]. Recent studies have found that NORC residents are older, more female, have lower incomes, live with more chronic conditions, and access home and primary care more often than non-NORC residents [43]. Together, these results suggest that there are more significant health needs than those of the general aging population; however, their communal living environment is particularly well-suited for leveraging. In this way, having a large concentration of older adults allows one to address technology needs in the same way health and social programming is delivered in NORCs [44], mainly through partnerships and collective efforts [45]. While we did not collect NORC participants' income levels, medical histories, or educational backgrounds, leveraging the collective power of a NORC environment to arrange, for example, technology training was mentioned during the World Café forum.

Despite the increased interest in equity-informed technology for older adults, limited research exists regarding NORC residents' adoption of technology that supports aging-in-place. Our findings align with previous interventional technology research (eg, sensory-based passive remote monitoring and smart speakers) implemented in NORC-like environments with lower-income and diverse older adults [46,47]. In these examples, access to a "cultural navigator" supported technology adoption through close relationships and information sessions [46]. However, without high-touch support for technology, its use among older residents may be limited to essential functions [47]. Similarly, NORC residents in another study expressed frustration and fear with using technology to access virtual programming during COVID-19 [48].

Strengths and Limitations

Our study uniquely positions NORCs and related collaborative relationships as one approach to consider when addressing the digital divide. Learnings from the NORC context could inform future digital health research, including design and implementation studies focusing more broadly on aging-in-place. The research agenda drawn from this study provides guiding questions that could help identify priority areas when considering equity-informed technology deployment and use in NORCs; they include viewing the opportunity to support aging in place with technology through various lenses. Finally, multiple data sources informed the agenda development, including a rapid review, a focus group, and a World Café. Despite these strengths, this study is not without limitations. For instance, the study data reflect a group of NORC residents and stakeholder participants. Our sample may not be generalized across other NORC settings, such as those in many rural and northern communities and across racial and ethnic minority and low-income older adults. In these settings, limited high-speed broadband internet continues to be problematic [49]. Likewise, since the target population in our study was NORC residents, certain strategies could be used to gather similar information from other population groups. As noted in the literature, partnering with community organizations, relevant clinicians, health service planners, and policy makers can leverage existing networks while establishing new relationships [28]. These networks not only serve to optimize participation among different community members, but they could also support ongoing initiatives catalyzed by the engagement. We also acknowledge that focus groups and World Café participants from NORCs had some degree of technology experience, which may indicate that we did not hear directly from those with minimal technical literacy. Therefore, a more targeted recruitment approach could have identified those significantly impacted by the digital divide. Another consideration is the potential influence that receiving the gift card could have had on participants' decisions regarding participation, for instance, being unduly incentivized to join the World Café. Finally, our final research agenda did not undergo a consensus exercise with our study participants. Future research could involve finalizing the agenda through a more iterative process, asking participants to rank questions on impact and importance.

Future Directions

Future research could tackle one or more of these questions by identifying conventional and nontraditional intersectoral partnerships between, for example, the public library system, academic health institutions, technology start-ups, the public education system, third-sector and charity organizations, and community-based service and housing providers. These new collaborative intersectoral models could help provide more accessible assistance to older adults within a communal environment who seek technology support [50]. Our study has begun to explore NORCs as unique environments that offer a transformative opportunity to address equity considerations in technology supporting aging in place.

Conclusions

Our findings and research agenda highlight critical areas for consideration, including leveraging partnerships, integrating public and private technology ecosystems, and designing technology with older users that evolves with the population's needs. Notably, by embedding principles of equity, inclusivity, and user-centered design, developers, researchers, and service providers can ensure that emerging technology serves diverse aging populations equitably and effectively.

Acknowledgments

The authors thank all the focus group and World Café participants for their invaluable contributions to the study. The Canadian Institutes of Health Research supported this work (award RN503880 - 498841).

Conflicts of Interest

None declared.

References

1. Jakobi J. Aging in Place Reflections from Canadians: Thea Franke Consulting; 2022.
2. Go C. Aging and Chronic Diseases: A Profile of Canadian Seniors: Public Health Agency of Canada Ottawa; 2020.
3. Bosch-Farré C, Malagón-Aguilera MC, Ballester-Ferrando D, et al. Healthy ageing in place: Enablers and barriers from the perspective of the elderly. A qualitative study. Int J Environ Res Public Health 2020 Sep 4;17(18):6451. [doi: 10.3390/ijerph17186451] [Medline: 32899744]
4. Iciaszczyk N, Arulnamby A, Brydges M, Marshall M, Cheng S, Feil C. Ageing in the Right Place: Supporting Older Canadians to Live Where They Want: National Institute on Ageing (NIA); 2022.
5. Golant SM. The quest for residential normalcy by older adults: Relocation but one pathway. J Aging Stud 2011 Aug;25(3):193-205. [doi: 10.1016/j.jaging.2011.03.003]
6. Golant SM. Residential normalcy and the enriched coping repertoires of successfully aging older adults. Gerontologist 2015 Feb;55(1):70-82. [doi: 10.1093/geront/gnu036] [Medline: 24840917]
7. Cheek P, Nikpour L, Nowlin HD. Aging well with smart technology. Nurs Adm Q 2005;29(4):329-338. [doi: 10.1097/00006216-200510000-00007] [Medline: 16260997]
8. E J, Xia B, Susilawati C, Chen Q, Wang X. An overview of Naturally Occurring Retirement Communities (NORCs) for ageing in place. Buildings 2022;12(5):519. [doi: 10.3390/buildings12050519]
9. National Institute on Ageing and NORC Innovation Centre. It's Time to Unleash the Power of Naturally Occurring Retirement Communities in Canada: National Institute on Ageing, Toronto Metropolitan University and NORC Innovation Centre, University Health Network; 2022.
10. DePaul VG, Parniak S, Nguyen P, et al. Identification and engagement of naturally occurring retirement communities to support healthy aging in Canada: a set of methods for replication. BMC Geriatr 2022 Apr 22;22(1):355. [doi: 10.1186/s12877-022-03045-z] [Medline: 35459126]
11. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
12. Liu L, Stroulia E, Nikolaidis I, Miguel-Cruz A, Rios Rincon A. Smart homes and home health monitoring technologies for older adults: a systematic review. Int J Med Inform 2016 Jul;91:44-59. [doi: 10.1016/j.ijmedinf.2016.04.007] [Medline: 27185508]
13. Marshall BL, Dalmer NK, Katz S, Loos E, López Gómez D, Peine A. Digitization of aging-in-place: an international comparison of the value-framing of new technologies. Societies (Basel) 2022;12(2):35. [doi: 10.3390/soc12020035]
14. Spacey J. Low tech vs high tech. Simplicable. 2016. URL: https://simplicable.com/new/low-tech-vs-high-tech [accessed 2025-07-07]
15. MacInnes J, Billings J, Dima AL, Farmer C, Nijpels G. Achieving person-centredness through technologies supporting integrated care for older people living at home: an integrative review. J Integr Care 2021 Jul 19;29(3):274-294. [doi: 10.1108/JICA-03-2021-0013]
16. Kim D, Bian H, Chang CK, Dong L, Margrett J. In-home monitoring technology for aging in place: scoping review. Interact J Med Res 2022 Sep 1;11(2):e39005. [doi: 10.2196/39005] [Medline: 36048502]
17. Grigorovich A, Fang ML, Sixsmith J, Kontos P. Defining and evaluating transdisciplinary research: implications for aging and technology. Disabil Rehabil Assist Technol 2019 Aug;14(6):533-542. [doi: 10.1080/17483107.2018.1496361] [Medline: 30318930]
18. Rubeis G, Fang ML, Sixsmith A. Equity in AgeTech for ageing well in technology-driven places: the role of social determinants in designing AI-based assistive technologies. Sci Eng Ethics 2022 Oct 27;28(6):49. [doi: 10.1007/s11948-022-00397-y] [Medline: 36301408]
19. Lee C. Technology and aging: the jigsaw puzzle of design, development and distribution. Nat Aging 2022 Dec;2(12):1077-1079. [doi: 10.1038/s43587-022-00325-6] [Medline: 37118534]
20. Greer V, Canham SL, Hong A, Agutter J, Garcia Zambrana I, Van Natter JM. Aging in place through the COVID-19 pandemic: perspectives from aging service providers. J Appl Gerontol 2023 Jul;42(7):1530-1540. [doi: 10.1177/07334648231159375] [Medline: 36856262]
21. van Dijk J. Digital divide research, achievements and shortcomings. Poetics (Amst) 2006 Aug;34(4-5):221-235. [doi: 10.1016/j.poetic.2006.05.004]
22. Andrey S, Masoodi MJ, Malli N, Dorkenoo S. Mapping Toronto's digital divide 2021. : Ryerson Leadership Lab and Brookfield Institute for Innovation + Entrepreneurship; 2023 Feb 1 URL: https://www.toronto.ca/wp-content/uploads/2021/07/95dc-Mapping-Torontos-Digital-Divide.pdf [accessed 2025-07-24]
23. Fang ML, Siden E, Korol A, Demestihas MA, Sixsmith J, Sixsmith A. A scoping review exploration of the intended and unintended consequences of eHealth on older people: a health equity impact assessment. Hum Technol 2018;14(3):297-323. [doi: 10.17011/ht/urn.201811224835]

24. Kokorelias KM, Nelson ML, Tang T, et al. Inclusion of older adults in digital health technologies to support hospital-to-home transitions: secondary analysis of a rapid review and equity-informed recommendations. JMIR Aging 2022 Apr 27;5(2):e35925. [doi: 10.2196/35925] [Medline: 35475971]
25. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: the significance of race/ethnicity. Gerontologist 2019 Jan 9;59(1):6-14. [doi: 10.1093/geront/gny138] [Medline: 30452660]
26. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G]
27. Boyko JA, Lavis JN, Abelson J, Dobbins M, Carter N. Deliberative dialogues as a mechanism for knowledge translation and exchange in health systems decision-making. Soc Sci Med 2012 Dec;75(11):1938-1945. [doi: 10.1016/j.socscimed.2012.06.016]
28. MacFarlane A, Galvin R, O'Sullivan M, et al. Participatory methods for research prioritization in primary care: an analysis of the World Café approach in Ireland and the USA. Fam Pract 2017 Jun 1;34(3):278-284. [doi: 10.1093/fampra/cmw104] [Medline: 27677298]
29. Monforte J, Netherway J, Smith B. The world café is an unmethod within co-produced research. Qual Res Psychol 2023 Jul 3;20(3):398-419. [doi: 10.1080/14780887.2023.2239728]
30. Löhr K, Weinhardt M, Sieber S. The "World Café" as a participatory method for collecting qualitative data. Int J Qual Methods 2020 Jan 1;19:1609406920916976. [doi: 10.1177/1609406920916976]
31. Khong L, Bulsara C, Hill KD, Hill AM. How older adults would like falls prevention information delivered: fresh insights from a World Café forum. Ageing Soc 2017 Jul;37(6):1179-1196. [doi: 10.1017/S0144686X16000192]
32. Street J, Barrie H, Eliott J, et al. Older adults' perspectives of smart technologies to support aging at home: insights from five World Café forums. Int J Environ Res Public Health 2022 Jun 25;19(13):7817. [doi: 10.3390/ijerph19137817] [Medline: 35805477]
33. Huynh T, Sava N, Hahn-Goldberg S, et al. Mobile on-site COVID-19 vaccination of naturally occurring retirement communities by neighbourhood risk in Toronto. Science Briefs - Ontario COVID-19 Science Advisory Table. 2021. URL: https://covid19-sciencetable.ca/sciencebrief/mobile-on-site-covid-19-vaccination-of-naturally-occurring-retirement-communities-by-neighbourhood-risk-in-toronto/ [accessed 2024-07-24]
34. Ollevier A, Aguiar G, Palomino M, Simpelaere IS. How can technology support ageing in place in healthy older adults? A systematic review. Public Health Rev 2020 Nov 23;41(1):26. [doi: 10.1186/s40985-020-00143-4] [Medline: 33292707]
35. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
36. Stokols D. Establishing and maintaining healthy environments. Toward a social ecology of health promotion. Am Psychol 1992 Jan;47(1):6-22. [doi: 10.1037//0003-066x.47.1.6] [Medline: 1539925]
37. Olaniyan A, Isiguzo C, Hawk M. The Socioecological Model as a framework for exploring factors influencing childhood immunization uptake in Lagos state, Nigeria. BMC Public Health 2021 May 5;21(1):867. [doi: 10.1186/s12889-021-10922-6] [Medline: 33952252]
38. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
39. Chan DYL, Lee SWH, Teh PL. Factors influencing technology use among low-income older adults: a systematic review. Heliyon 2023 Sep;9(9):e20111. [doi: 10.1016/j.heliyon.2023.e20111] [Medline: 37809586]
40. Suntai Z, Beltran SJ. The intersectional impact of race/ethnicity and sex on access to technology among older adults. Gerontologist 2023 Aug 24;63(7):1162-1171. [doi: 10.1093/geront/gnac178] [Medline: 36477498]
41. Raihan MMH, Subroto S, Chowdhury N, Koch K, Ruttan E, Turin TC. Dimensions and barriers for digital (in)equity and digital divide: a systematic integrative review. Digit Transform Soc 2025 May 6;4(2):111-127. [doi: 10.1108/DTS-04-2024-0054]
42. Sheth MS, Rochon PA, Altaf A, et al. Factors associated with transition to a nursing home in older adults living in naturally occurring retirement communities. J Am Geriatr Soc 2025 Jan;73(1):88-100. [doi: 10.1111/jgs.19038] [Medline: 38888381]
43. Savage RD, Huynh T, Hahn-Goldberg S, et al. A portrait of older adults in naturally occurring retirement communities in Ontario, Canada: a population-based study. J Am Geriatr Soc 2025 Jan;73(1):74-87. [doi: 10.1111/jgs.19278] [Medline: 39618101]
44. Mills CM, Parniak S, Hand C, et al. The impact of a naturally occurring retirement community supportive services program on older adult participants' social networks: semistructured interview study. JMIR Aging 2022 Nov 21;5(4):e37617. [doi: 10.2196/37617] [Medline: 36409533]
45. Parniak S, DePaul VG, Frymire C, DePaul S, Donnelly C. Naturally occurring retirement communities: scoping review. JMIR Aging 2022 Apr 14;5(2):e34577. [doi: 10.2196/34577] [Medline: 35436204]
46. Berridge C, Chan KT, Choi Y. Sensor-based passive remote monitoring and discordant values: qualitative study of the experiences of low-income immigrant elders in the United States. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11516. [doi: 10.2196/11516] [Medline: 30907741]

47. Chung J, Bleich M, Wheeler DC, et al. Attitudes and perceptions toward voice-operated smart speakers among low-income senior housing residents: comparison of pre- and post-installation surveys. Gerontol Geriatr Med 2021;7:23337214211005869. [doi: 10.1177/23337214211005869] [Medline: 35047655]
48. Fernandes K, Hand C, Laliberte Rudman D, et al. Being and doing together in a naturally occurring retirement community: pandemic experiences of older adults. Can J Aging 2025 Mar;44(1):115-125. [doi: 10.1017/S0714980824000345] [Medline: 39610356]
49. Freeman S, Marston HR, Ross C, et al. Progress towards enhanced access and use of technology during the COVID-19 pandemic: a need to be mindful of the continued digital divide for many rural and northern communities. In: Gates J, editor. Healthc Manage Forum 2022 Sep;35(5):286-290. [doi: 10.1177/08404704221108314] [Medline: 35855623]
50. Golant SM. Commentary: Irrational exuberance for the aging in place of vulnerable low-income older homeowners. J Aging Soc Policy 2008;20(4):379-397. [doi: 10.1080/08959420802131437] [Medline: 19042553]

Abbreviations

NORC: naturally occurring retirement community
SEM: Social Ecological Model

Edited by AF Abdulai; submitted 09.01.25; peer-reviewed by B Davis, J Scott, N Novik; revised version received 23.05.25; accepted 20.06.25; published 30.07.25.

Please cite as:
Saragosa M, Abejirinde IOO, MacEachern E, Nelson MLA, Kokorelias KM, Bharmal S, Ludwig Prout B, Mohamed M
Exploring Technology Supporting Aging-in-Place Using an Equity Lens Through Focus Groups and World Café–Informed Research Agenda: Qualitative Study
JMIR Aging 2025;8:e71093
URL: https://aging.jmir.org/2025/1/e71093
doi:10.2196/71093

Culturally Adapted STAR-Caregivers Virtual Training and Follow-Up for Latino Caregivers of People Living With Dementia: Single-Arm Pre-Post Mixed Methods Study

Miguel Angel Mariscal, MPH; Celeste Garcia, MPH; Lily Zavala, BS; Magaly Ramirez, PhD

Department of Health Systems and Population Health, School of Public Health, University of Washington, Hans Rosling Center for Population Health, 3980 15th Avenue NE, Fourth Floor, Box 351621, Seattle, WA, United States

Corresponding Author:
Magaly Ramirez, PhD
Department of Health Systems and Population Health, School of Public Health, University of Washington, Hans Rosling Center for Population Health, 3980 15th Avenue NE, Fourth Floor, Box 351621, Seattle, WA, United States

Abstract

Background: Latino caregivers are at an increased risk of negative health outcomes due to the responsibilities of caring for someone with dementia. Although interventions exist to address caregiver burden, they often do not meet the cultural needs of Latino caregivers.

Objective: This study aimed to pilot test the cultural adaptation of the STAR-Caregivers Virtual Training and Follow-Up (STAR-VTF) intervention. The intervention is an evidence-based training program designed to teach family caregivers strategies to manage behavioral and psychological symptoms of dementia (BPSD). Our research team has conducted past studies to identify and perform culturally relevant adaptations to the training modules of STAR-VTF, and this study aimed to pilot these culturally adapted modules with a sample of Latino caregivers.

Methods: Data on feasibility, usability, and acceptability were collected from a pilot test in which Latino caregivers (n=16) used the training modules of the STAR-VTF intervention over a 7-week period. Participants completed usability surveys following the completion of each module, and acceptability was assessed through semistructured interviews (n=14) postintervention. Preliminary outcome measures were also collected, and a descriptive analysis was conducted. The primary outcomes were the Revised Memory and Behavior Problem Checklist (RMBPC) and the Preparedness for Caregiving Scale.

Results: The pilot study results suggest that it is feasible to deliver the culturally adapted STAR-VTF intervention to Latino caregivers, with 94% (15/16) of participants maintaining enrollment through intervention completion. The intervention's usability was found to be "good" based on an average System Usability Score of 76.7 out of 100 across all training modules. Caregivers were generally satisfied with the training modules. In addition, preliminary outcome results demonstrated a trend of decreased BPSD pre- versus postintervention (RMBPC subscale score: 28.24 to 21.34). Findings also demonstrated decreased caregiver reaction to BPSD pre- versus postintervention (RMBPC subscale score: 40.40 to 37.21) and increased caregiver preparedness based on pre- and postintervention (Preparedness Caregiving Scale score: 1.98 to 2.43).

Conclusions: The pilot study demonstrated that the culturally adapted STAR-VTF intervention is feasible and perceived as easy to use by a small sample of Latino caregivers. We aim to refine the cultural adaptations of the STAR-VTF intervention further based on feedback from study participants. Future studies are necessary to test the efficacy of the intervention and support the broad dissemination of the culturally adapted intervention.

(JMIR Aging 2025;8:e66053) doi:10.2196/66053

KEYWORDS

dementia; gerontology; geriatric; older person; aging; Latina; Latino; caregiver; symptom management; digital technology; digital intervention; digital health application; pilot study

Introduction

In the United States, Latino populations are disproportionately impacted by Alzheimer disease and related dementias (ADRD). Compared to White adults, Latinos are 1.5 times more likely to have ADRD [1]. By 2060, the number of Latinos with ADRD is expected to grow by 832%, which is the steepest increase in ADRD compared to any other racial or ethnic group, and projections show that by 2060, a total of 3.5 million Latinos will be living with dementia [2,3]. Latinos with dementia primarily rely on family members for their care due to the high cost of formal dementia care, legal status, barriers to accessing dementia care services, language preferences, and cultural preferences [4,5].

Latino caregivers comprise 21% of the estimated 40 million family caregivers in the United States [1]. Caregiving for a person living with dementia can be particularly challenging as ADRD can cause changes in personality and behavior, losses in judgment, orientation, and the ability to understand and communicate effectively [6]. Managing dementia symptoms requires intense supervision and physical support, and compared to other populations, Latino families provide more intense caregiving in terms of level of care and hours [1]. Caregiving for a person living with dementia also places Latino caregivers at increased risk of negative health implications, including high levels of stress, depression symptoms, less engagement in physical activity, sleeping problems, and social isolation [4,7]. Although evidence-based caregiver interventions exist to help address the negative impacts of caregiver burden, they often do not meet the unique cultural needs of Latino families [8]. Due to the negative health implications of dementia caregiving, it is crucial to support Latino caregivers through culturally adapted, evidence-based interventions.

STAR Caregivers (STAR-C) is an evidence-based systematic training program designed to teach family caregivers strategies to reduce behavioral distress and manage behavioral and psychological symptoms of dementia [8]. Caregivers are taught effective communication strategies, activator-behavior-consequence (ABC) behavioral problem-solving strategies, and how to identify and use pleasant activities to reduce behavioral disturbances [9]. STAR-C has been demonstrated to significantly reduce caregiver burden and depression and improve the quality of life of persons with ADRD [10]. Since the original randomized controlled trial of STAR-C in 2005 [11], the program has been adopted in various settings to meet caregivers' needs. One of the early adoptions of the STAR-C program in a real-world setting was conducted by the Oregon Department of Human Services-State Unit on Aging and Area Agencies of Aging, which served both rural and urban locations in Oregon [9,12]. In this implementation of STAR-C, the training was provided at home by trained professional health consultants to family caregivers. The program has since been adapted to teach caregivers the STAR-C lessons remotely. Tele-STAR, the telehealth-based adaptation of STAR-C, was created to increase access to anyone with a computer and internet connection and to reach rural populations [13]. Tele-STAR used videoconferencing to connect nurse consultants with family caregivers to guide them through the STAR-C lesson plans. The tele-STAR study found that with the transition from in-person training to remote, the family caregiver burden was reduced, and it retained good program and treatment fidelity to STAR-C [13]. Videoconferencing technology has also been used in the Illinois-based CRIS' Memory Care Program, an adapted STAR-C intervention developed by CRIS Healthy Aging, a nonprofit organization in East Central Illinois that serves older adults primarily in rural communities and of lower income [14]. The latest adaptation of STAR-C in a clinical setting has been STAR-C Virtual Training and Follow-up (STAR-VTF), which was tested at Kaiser Permanente Washington [15]. STAR-VTF is a 6-to-8-week program comprised of six training modules that caregivers complete asynchronously. It also includes 30-minute weekly check-ins with a program coach. Support from program coaches is provided as needed for up to 6 months via secure messaging in the Kaiser Permanente Washington patient portal.

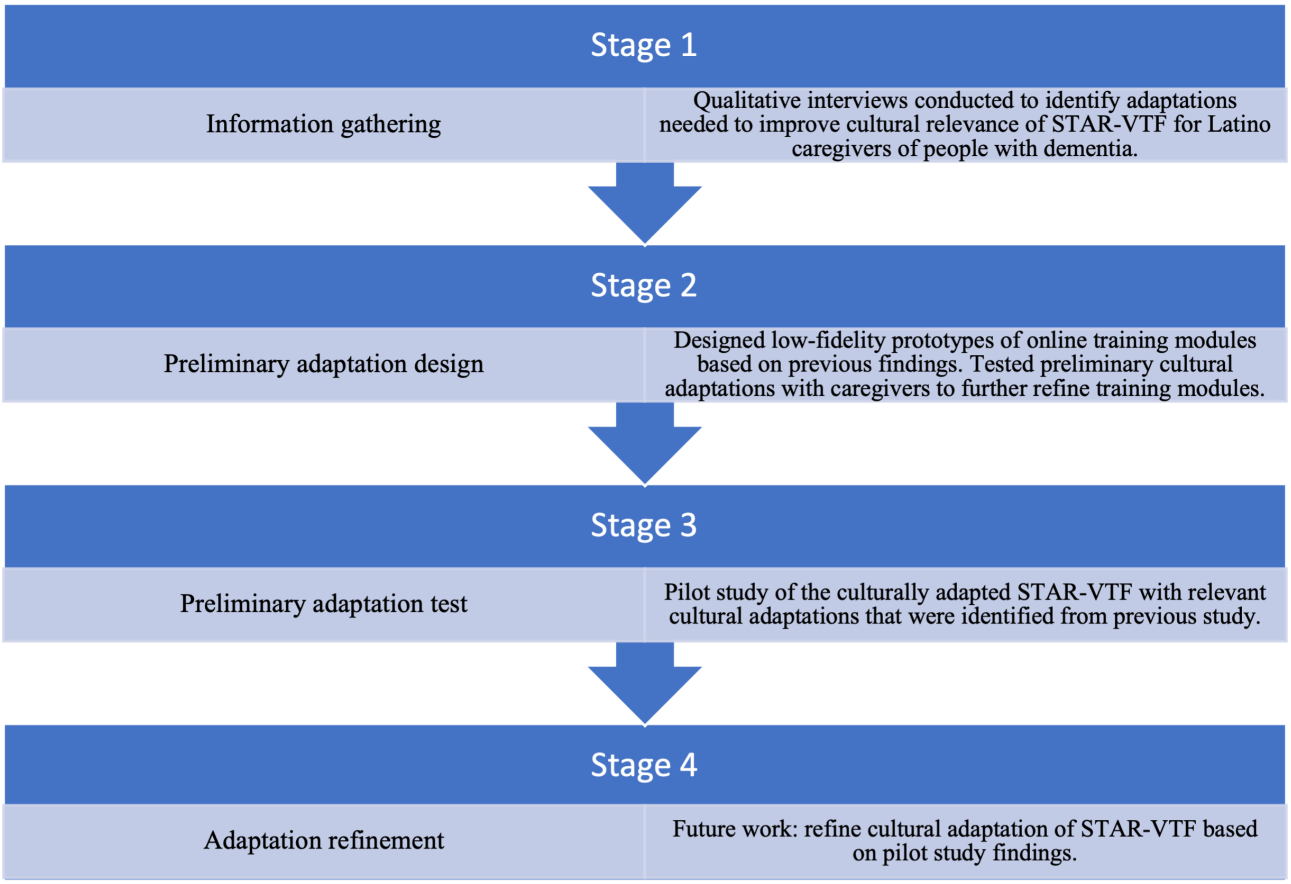
Although STAR-C has been adapted in a variety of ways to expand access to caregivers, it has yet to be tested in large populations of Latino caregivers. Recently, our research team adapted STAR-VTF to improve its cultural relevance and linguistic appropriateness for Latino caregivers [8,15]. We conducted an 8-week pilot study of the culturally adapted training modules from the STAR-VTF intervention. This study reports findings on the feasibility, usability, acceptability, and preliminary outcomes of the pilot study with Latino caregivers. The findings will help improve future iterations of STAR-VTF for Latino caregivers, and more generally, advance our understanding of modifying evidence-based interventions to better serve Latino caregivers of persons with dementia.

Methods

Design

The framework for the cultural adaptation of evidence-based interventions developed by Barrera and Castro [16] guided the methods of this study. The framework includes the following stages in culturally adapting an evidence-based intervention: (1) gathering information to identify ideas about needed adaptations, (2) conducting preliminary adaptations based on these ideas, (3) conducting pilot studies of the preliminary adaptations, and (4) refining adaptations based on results from pilot studies. Our research team has completed the first two stages [8,15]. This study presents findings in the stage 3 process of pilot-testing cultural adaptations (Figure 1).

Figure 1. STAR-VTF cultural adaptation process. STAR-VTF: STAR-Caregivers Virtual Training and Follow-Up.



This single-arm pilot study tested the feasibility, usability, and acceptability of the culturally adapted STAR-VTF training modules. We also trialed data collection procedures and collected preliminary data on persons living with dementia and caregiver outcomes pre- and postintervention to obtain descriptive estimates. Quantitative and qualitative methods were used in this pre- and posttest study.

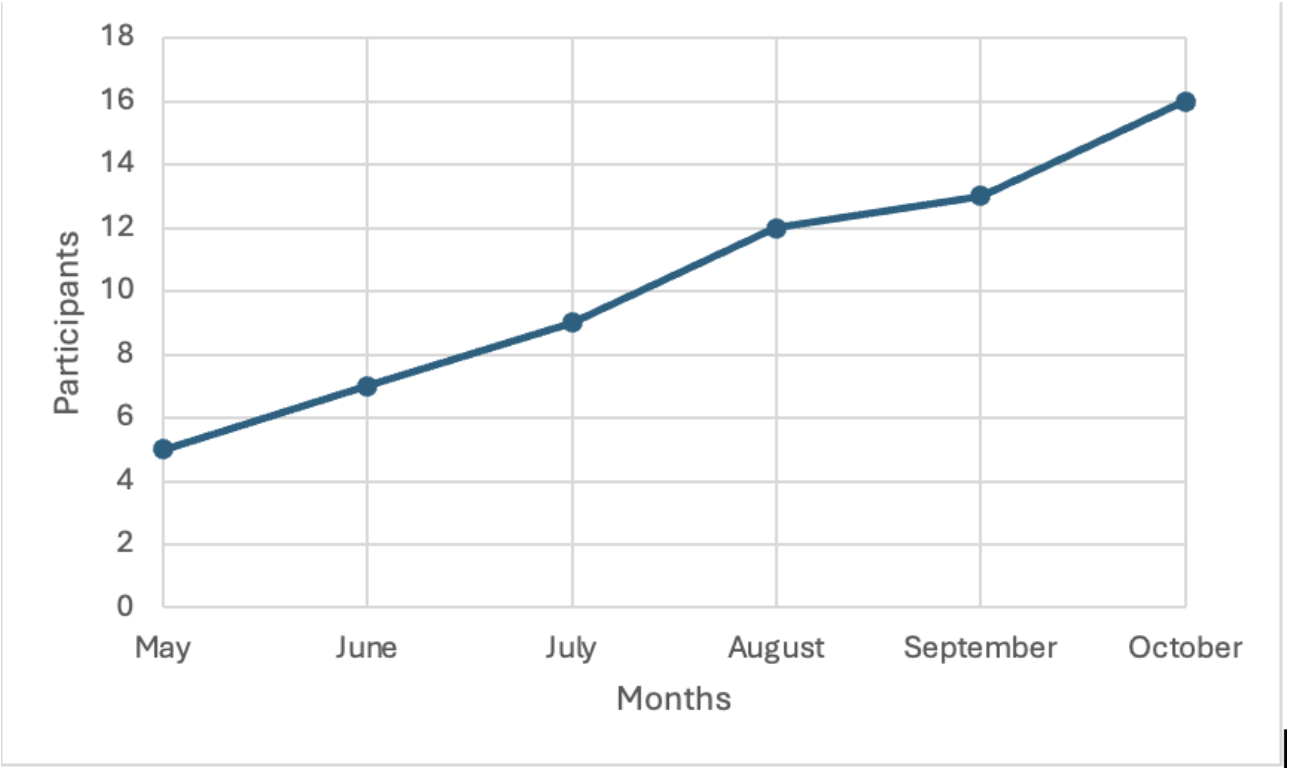
Ethical Considerations

The Advarra Institutional Review Board approved this study (Protocol: National Institute on Aging - FY20_Pilot10_Ramirez). If interested, participants were screened for eligibility, and digital or written informed consent was obtained. Consent was collected from participants via a consent form that was provided in English and Spanish, depending on their preference. Translation of the consent form was done by a certified Spanish translator. Participants were provided a US $40 gift card for their participation. All study data were deidentified and stored on secure, password-protected cloud storage.

Participant Recruitment

Potential study participants (ie, Latino caregivers) were identified using the UW Medicine electronic health record, UW Alzheimer's Disease Research Center Registry, Alzheimer's Prevention Registry maintained by the Banner Alzheimer's Institute, and through the distribution of recruitment flyers through professional networks. Potential study participants were contacted by bilingual staff via phone or email to explain the study and answer their questions about study participation. To be eligible to participate, participants had to identify as Hispanic or Latino, be 18 years of age or older, be taking care of a family member or friend with dementia, live with or within 5 miles of the person living with dementia, spend at least 8 hours a week with them, and indicate that the person they care for is experiencing at least three symptoms related to dementia occurring at least three times in the past week. Figure 2 provides additional information on the number of participants who were enrolled over time for this pilot study.

Figure 2. Enrollment timeline.



Intervention: Culturally Adapted STAR-VTF Training Modules

While the full STAR-VTF intervention consists of weekly training modules and weekly 30-minute phone check-ins with a coach [15], this study only pilot-tested the culturally adapted modules. The adaptations used in this pilot study were based on findings from previous qualitative studies [8,15]. These studies interviewed Latino caregivers of persons with dementia, as well as health care and social service providers of older Latinos, to improve the cultural relevance of STAR-VTF. Based on the findings of prior studies, adaptations included revising language viewed as problematic, expanding content to enhance understanding of dementia, and adding cultural examples that reflect family involvement in caring for relatives with dementia and multigenerational living [8,15]. Additional refinements made to the STAR-VTF training modules, after testing preliminary modules with Latino caregivers, included adding empathetic messaging that highlights the importance of viewing the world from the perspective of the person living with dementia, incorporating additional problem-solving examples to demonstrate diverse challenges, and emphasizing caregiver self-care, to ensure caregiver's mental and emotional health is prioritized [8].

For 7 weeks, caregivers completed training modules asynchronously. They accessed the modules via email or SMS text message. This pilot study was focused on testing the content of the training modules; therefore, caregivers did not receive the coaching component of the STAR-VTF intervention. There were 7 training modules in total. Caregivers were instructed to complete one module per week. The content of the modules is described in Textbox 1.

Textbox 1. Content of the STAR-Caregivers Virtual Training and Follow-Up training modules.

Module 1

- Provides an understanding and overview of dementia.

Module 2

- Introduces caregivers to the behavioral treatment of dementia, realistic expectations, and effective communication.

Module 3

- Covers the antecedents, behaviors, consequences (ABC) approach to problem-solving, including rationale and development of an ABC plan. Hypothetical scenarios of how caregivers successfully used the ABC problem-solving to address behaviors (eg, wandering) are included.

Module 4

- Instructs caregivers to review the ABC plan and revise it as needed.

Module 5

- Covers pleasant events and managing negative thinking.

Module 6

- Instructs caregivers to review the ABC plan, pleasant activities schedule, and to revise as needed.

Module 7

- Covers coping with caregiving and maintaining the use of caregiving strategies.

The length of the training modules ranged between 7 and 16 minutes each. The modules used text, pictures, and illustrations with a voiceover presentation. Caregivers received the training modules in their preferred language (English or Spanish). They also received a printed or electronic workbook to accompany the lessons (the format depended on their preference). The workbook contained additional educational materials, as well as worksheets for caregivers to write down their ABC plans. Intervention materials, including training modules, voiceovers for modules, and workbooks, were translated into Spanish by a certified Spanish translator. Study participants were enrolled in the study for approximately 8 weeks.

Measures

Feasibility

We assessed feasibility based on National Institutes of Health guidance on appropriate feasibility measures for pilot studies [17]. Feasibility measures included details of the success of procedures to recruit eligible participants, retention, and intervention completion.

Usability

To assess usability, we collected data on the System Usability Scale [18] via a REDCap (Research Electronic Data Capture; Vanderbilt University) survey each time a caregiver completed a training module. With the System Usability Scale survey, caregivers rated how easy the training modules were to use. Each item was rated on a 5-point scale ranging from 0=strongly disagree to 5=strongly agree.

Acceptability

Assessment of acceptability was accomplished through semistructured qualitative interviews with participants after they completed 7 weeks of the STAR-VTF training modules. Interview questions explored the participant's experience with the program and solicited feedback on recommendations for further improving the program.

Person Living With Dementia and Caregiver Outcomes

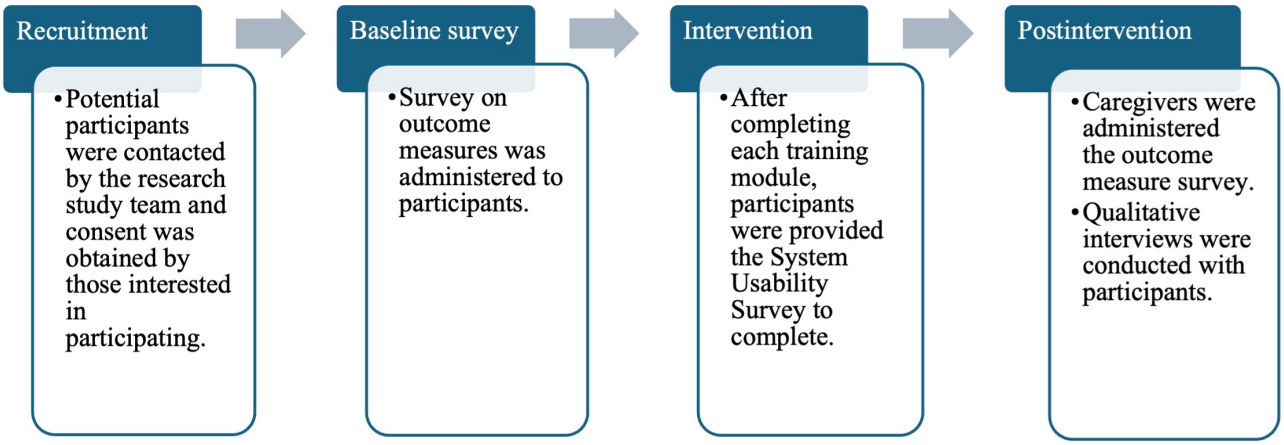
The primary outcomes were the Revised Memory and Behavior Problem Checklist (RMBPC) [19,20] and the Preparedness for Caregiving Scale [21]. The outcomes were collected at baseline and 8 weeks postenrollment via a REDCap survey. For the RMBPC, caregivers rated memory, depression, and disruptive behavior problems in the person living with dementia. The instrument contains 24 items (7 memory-related, 8 depressive, and 9 disruptive) that assess problem behaviors and are rated for frequency of occurrence during the past week and caregiver reaction, which refers to the extent to which caregivers were distressed by the problem behaviors. The Preparedness for Caregiving Scale was used for caregivers to rate how prepared they are for various aspects of caregiving. The instrument contains 8 items that ask caregivers how well prepared they believe they are to provide physical care and emotional support, deal with the stress of caregiving, and set up in-home support services. Each item is rated on a 5-point scale ranging from 0=not at all prepared to 5=very well prepared.

Data Collection

Timeline of Data Collection

The data for this study were collected pre- and postintervention and throughout the intervention period. For reference to the data collection timeline, see Figure 3.

Figure 3. Time of data collection.



Quantitative Data Collection

We collected outcomes data on the RMBPC, Preparedness for Caregiving Scale, and System Usability Scale via REDCap surveys in Spanish and English. The Spanish versions of these scales have been validated and found reliable [20,22,23]. Research assistants on the project delivered a link to the surveys via email or SMS text message to caregivers. These surveys were provided to participants at baseline and postcompletion of the training modules.

Qualitative Data Collection

All participants who reached the end of the study were invited to participate in a semistructured interview conducted in Spanish or English. All interviews with participants were conducted over the phone with bilingual research team members (CG and LZ). The interview guide included questions about their opinions on the program training components, suggestions for improving course content and design, and the quality of the Spanish translation of training modules (Multimedia Appendix 1). The average length of interviews was approximately 30 minutes. Interviews with participants were audio recorded, and the transcription of the interviews was done by a professional transcription service.

Quantitative Data Analysis

We analyzed quantitative data using descriptive statistics. Based on National Institutes of Health guidance, we did not conduct outcome analysis for effect size estimates since an effect size estimated from a pilot study is inherently unstable and would not provide helpful estimation for power calculations [17].

In the descriptive analysis, we only included study participants who completed the outcome assessments at baseline and 8 weeks. If study participants had missing values in their responses (either because they skipped an item or responded with “Don’t Know/Not Applicable”), then we imputed their missing values. The missing data were imputed by taking the highest possible score that participants could obtain if they answered all items and then multiplying that score by the quotient of the participant’s actual score divided by the highest possible score they could have received from the items they answered.

Qualitative Data Analysis

To analyze qualitative data, we used qualitative analysis methods outlined by Miles et al [24]. Interview transcripts were coded and analyzed in their original language (English or Spanish) using Dedoose (version 9.2.007), a cloud app for managing, analyzing, and presenting qualitative and mixed method research data. First, bilingual team members (MAM, CG, and MR) read the transcripts and wrote brief notes of what was in the data. Next, the first (MAM) and second (CG) authors independently coded transcripts using inductive codes. They convened weekly with the senior author (MR) to discuss the application of codes, refine the codebook consisting of inductive codes, and reach a consensus on coding discrepancies. After coding three transcripts, the codebook was comprehensive, and a few additional codes needed to be added. The first author individually coded the remaining transcripts, with guidance from the senior author as needed. Descriptive and in vivo codes were used in the coding process.

The first author identified preliminary themes based on the inductive codes by collating and reviewing excerpts of each code. During this process, relationships among codes and relationships within and among themes were assessed. A collection of relevant excerpts for each theme was pulled from the data while reviewing coded excerpts. Recommended areas of assessment in the third stage of Barrera and Castro’s [16] framework on cultural adaptations of evidence-based interventions, such as satisfaction with intervention elements and suggestions for improvement, were used to help guide theme creation. The preliminary themes were iteratively refined with assistance from the second and third authors (CG and MR) to ensure that subthemes within the themes were relevant to the parent theme and that the final themes were distinguishable.

Results

Participant Characteristics

In total, 16 participants enrolled in the pilot study of the culturally adapted STAR-VTF training modules. Table 1 provides demographic characteristics of the study participants. Demographic characteristics were collected at baseline.

Table . Participant characteristics.

Characteristics	Study participants (n=16)
Age, mean (SD)	51.2 (13.2)
Gender, n (%)	
Man	4 (25)
Woman	12 (75)
Hispanic or Latino origin, n (%)	
Mexican, Mexican American, Chicano or Chicana	10 (62)
Puerto Rican	1 (6)
Other	5 (31)
Race, n (%)	
Black or African American	1 (6)
White	11 (69)
Other	3 (19)
Prefer not to answer	1 (6)
Highest level of education completed, n (%)	
Primary level education	1 (6)
Secondary level education or GED[a] or equivalent	2 (12)
Some college	2 (12)
Associate's degree	2 (12)
Trade or vocational training	1 (6)
Bachelor's degree or higher	8 (50)
Household income before taxes (US $), n (%)	
Less than 15,000	1 (6)
Between 15,000 and 19,999	1 (6)
Between 20,000 and 24,999	1 (6)
Between 25,000 and 34,999	1 (6)
Between 35,000 and 49,999	2 (12)
Between 50,000 and 74,999	2 (12)
100,000 or above	6 (37)
Prefer not to answer	2 (12)
Current occupational status, n (%)	
Employed	9 (56)
Unemployed	3 (19)
Stay-at-home caregiver	2 (12)
Retired	2 (12)
Current marital status, n (%)	
Married	12 (75)
Divorced	1 (6)
Separated	1 (6)
Never married	2 (12)
Primary language spoken at home, n (%)	
English	8 (50)
Spanish	8 (50)

Characteristics	Study participants (n=16)
Received STAR-VTF intervention materials in Spanish, n (%)	7 (44)
Caregiver's health insurance plan, n (%)	
Insurance through employer	8 (50)
Insurance purchased directly	1 (6)
Medicare	3 (19)
Medicaid	1 (6)
No Insurance	3 (19)
Person living with dementia health insurance plan, n (%)	
Insurance purchased directly	1 (6)
Medicare	9 (56)
Medicaid	5 (31)
TRICARE or military health care	1 (6)
Persons living with caregiver, n (%)[b]	
Spouse or partner	12 (75)
Children or grandchildren older than 18 years of age	5 (31)
Children or grandchildren younger than 18 years of age	4 (25)
Relatives (besides spouse or partner and children or grandchildren)	7 (44)
Other	1 (6)
Household size, mean (SD)	3.4 (1.5)
Devices owned, n (%)[b]	
Smartphone	15 (94)
Tablet	11 (69)
Laptop	11 (69)
Computer	7 (44)
Internet access, n (%)	
Cellular data plan	5 (31)
Broadband internet	10 (62)
Satellite internet	1 (6)
Caregiver's relationship to the person living with dementia, n (%)	
Spouse or partner	3 (19)
Adult child	9 (56)
Other close family member	3 (19)
Friend or other nonrelative	1 (6)
Gender of person living with dementia, n (%)	
Man	4 (25)
Woman	12 (75)
Lives with person living with dementia, n (%)	10 (62)
Years in the caregiving role, mean (SD)	4.3 (3.28)
Caregiving hours per week, mean (SD)	60.2 (50.9)

[a]GED: General Educational Development.

[b]Column percent may not add up to 100% because categories are not mutually exclusive.

Feasibility

We achieved satisfactory feasibility based on recruitment, retention, and intervention completion measures. We recruited 16 study participants. The initial recruitment goal was to recruit 20 participants, and we reached 80% of our goal (16/20). Of the participants recruited, 94% (15/16) of participants were retained for all 8 weeks of the pilot study. The proportion of participants who completed all 7 training modules and participated in all study procedures (qualitative exit interviews, pre- and postoutcome assessments, and weekly usability surveys) was 75% (12/16).

Usability

Usability was assessed for each training module. Average scores and SDs for each module were calculated (Table 2). The average System Usability Score across all 7 modules is 76.7 out of 100, which is considered "good" usability, indicating that caregivers perceive the STAR-VTF training modules as easy to use [24,25].

Table . System Usability Scores.

Week	System Usability Score, mean (SD)
1 (n=13)	77.8 (10.8)
2 (n=12)	77.7 (12.9)
3 (n=10)	74.6 (19.2)
4 (n=13)	73.7 (16.3)
5 (n=12)	75.8 (14.3)
6 (n=9)	79.2 (14.5)
7 (n=11)	78.0 (13.7)

Acceptability

The results of the semistructured interviews provided information about the cultural adaptations of the STAR-VTF training modules. The qualitative results also provide a further understanding of what changes are needed to revise the cultural adaptations further.

Theme 1: Latino Caregivers Were Satisfied With the STAR-C Intervention Content and Design Elements of the Training Modules and Accompanying Workbook

Overview

Latino caregivers were generally satisfied with the content covered in the STAR-C intervention, including components on dementia education, ABC problem-solving, and caregiver support strategies. They also expressed satisfaction with the design of the training modules and the accompanying workbook.

Dementia Education Component

Latino caregivers found the dementia education information beneficial in understanding their relatives' dementia condition and how it can progress. There was also a consensus among most caregivers that they liked the empathetic messaging embedded within the dementia education component. Caregivers felt that it was helpful to be reminded that their family member has a condition of the brain and that some of the things their family member says or does are not on purpose, and to be encouraged to have patience and understanding. A caregiver providing care to their father said the following about the messages reminding caregivers to be empathetic:

> *And I felt like it was an important reminder because I guess just like with anything else, even as a parent of young children, I get caught up in the day-to-day and the getting things done, and there's deadlines, and there's work and home life and all sorts of things. And so it is nice to have that reminder that it's not just ... not to take things personally and not to be so quick to react. But just a gentle reminder, there may be another reason why my dad is saying something or not doing something. And yeah, just not to take it personally, so that was a helpful reminder.* [ID 11]

ABC Problem-Solving Component

Many Latino caregivers expressed that the ABC problem-solving strategy of the training was helpful and found the examples used to demonstrate how to implement the strategy to be relatable to their own caregiving experience. One of the training modules demonstrated a hypothetical caregiver using the ABC problem-solving strategy to address a person living with dementia wandering. When asked about their thoughts on this example, one of the caregivers stated:

> *So I thought it was very relatable. And then sort of just kind of showcasing what happened, I think, made it relatable in the sense that you can sort of put yourself in that situation. And then for me, I was like, "Okay, yeah. I could see how this happens, and then what can you do to utilize those ABCs to then basically make it useful for your own personal experience in a similar situation."* [ID 14]

Latino caregivers also shared how the ABC problem-solving strategy was successful in helping them manage their relative's behavioral and psychological symptoms of dementia and provided anecdotes of their experience applying the ABC problem-solving strategy. For example, one of the caregivers described how they used the strategy to address their mother's experience of washing dishes. The caregiver said:

> *So my mom would get really upset and would complain of dizziness around the same time every day. And I thought that was odd. And then when I was doing this and I took the ABC, I realized that the*

reason my mom was getting upset was because she would wash the dishes at the same time every day at, and sometimes I would go by and tell her, "Oh, turn the water higher," or, "Oh, put more soap." And I think me trying to coach her was upsetting her because it made her feel like she didn't know what she was doing. So she would go, and she would say she's dizzy, and she would start to cry. And then because I noticed the timing, what happened before and after, I stopped doing that. I just let her wash the dishes. [ID 27]

Caregiver Support Strategies Component

Latino caregivers overwhelmingly related to the caregiver support strategies, encouraging caregivers to practice self-care. Caregivers expressed the importance of caring for themselves to care for others. They also acknowledged the importance of taking time for themselves. For example, a Spanish-speaking caregiver providing care to his wife explained the reasons why it was important for caregivers to prioritize their own self-care.

It's good that you made us see what we need, also we need to take care of ourselves to be able to provide care. In my case, my wife needed me to be good so I could take care of her well. One needs to sleep well and eat well to give the best help to the sick. (translated quote) [ID 06]

Design of Training Modules and Accompanying Workbook

Latino caregivers were satisfied with the design of the training modules and the accompanying workbook. Some caregivers said they liked the pacing of the training modules and the images used. One caregiver shared the following about the training modules:

I like the pace of it, and I liked the way that it was almost like going through a PowerPoint presentation. I liked that there were slides that kind of bulleted the info that I could either read along to or that it included images that were ... let me see. Images of people of color and not just an elderly White family. [ID 11]

Some of the caregivers described the workbook as a helpful resource to refer to information and strategies that had been covered in the training modules. For example, one of the caregivers planned to continue referencing the workbook after the research study ended. They stated:

So I think for me it's [the workbook] going to be a helpful guide now that I don't have access to the videos to go through it and remind me of things and I can share it with my other family members who are also trying to help with caring with my mom. So I feel like the workbook is kind of the legacy of those classes. [ID 27]

Theme 2: Suggestions for Improving Program Content and Making Training More Engaging

Overview

Although Latino caregivers were generally satisfied with the training modules, some caregivers shared suggestions that they felt would enhance the intervention. The areas for improvement that Latino caregivers suggested are the inclusion of information on additional resources and improving the interactivity of training modules to make videos more engaging.

Including Information on Additional Resources

Latino caregivers expressed wanting additional information on resources outside of the program that they can access to help support people living with dementia. Although some resources were presented in the training modules, caregivers shared they would like more resources on dementia care and community organizations serving people living with dementia and their families. For example, one caregiver shared:

Maybe what might be helpful is if you had ... now this would be just ... I guess you'd have to make it unique for every state, but maybe give them some resources of where they might be able to reach out to. There's different Medicaid programs that might be eligible for. There might be local dementia associations that they could contact, those types of things, right? So resources that people might be able to reach out to, especially ones that offer services in Spanish. [ID 17]

Improve Interactivity to Make Training Modules More Engaging

Some Latino caregivers provided suggestions for improving the interactivity of training modules to make learning program content more engaging. A suggestion to improve the interactivity of program content was to include knowledge checks or short quizzes in which caregivers can test their knowledge of the material covered. As one caregiver shared:

Having something that's a little more interactive where it's not just a video, but almost like a course where you watch a video and then ... having both options like an actual paper workbook or having the option to select multiple choice or just to keep the viewer engaged in the video. So even if it's like, "Here's a summary. Which of these apply?" Select A, B, C, or D, all of the above, or something like that, just so that it really just keeps the viewer engaged. [ID 11]

Latino caregivers also felt there were areas in which the training modules could be improved visually to make them more engaging. A suggestion that was shared by caregivers was to include video reenactments to demonstrate how strategies learned can be used in real-world situations. For example, one caregiver shared they liked the ABC strategy examples in the training module but mentioned that if they were recreated in a video depicting a role-play of the scenario, it would be more engaging. As the caregiver explains:

> *I feel like an actual video example of the dialogue would go a lot further than the pictures with the text-next-to-them pictures. If I remember correctly, that's what they were. They were pictures of the humans, of the woman and the person, her caregiver, and then there was just dialogue text next to them. It made sense. It just seems very dry. It seemed very dry. And I didn't really feel like I was paying attention as much as I should have in those moments.* [ID 24]

Some caregivers also suggested the importance of including testimonials from fellow caregivers of patients with dementia about their experiences with caregiving and how they applied the STAR-VTF training. One caregiver shared how incorporating caregiver testimonials can help demonstrate how to implement strategies covered in the training, such as the importance of caregiver self-care:

> *I think probably maybe interviewing some caregivers and having them share their journey from not knowing what's going on with your mother ... she's crazy ... to understanding what's happening and the realization that it is an illness and examples of how they continued on with taking care of their health. It's very stressful. It's very stressful. I mean, it creates a physical response in your own body. And so I think it's just the affirmation from others that what they've done to manage their own time and their own feelings and examples like that and reassuring us that it's okay. It's okay to leave them and do something fun like go to a movie with your husband instead of feeling guilty about it.* [ID 28]

Theme 3: Caregivers Share Dementia Care Information With Others and Suggest Outreach Strategies

Overview

Latino caregivers shared how they have increased awareness of dementia care strategies by sharing the information they learned in the training modules with relatives. Caregivers generally shared program content by word of mouth or by sharing the link to the training modules that were provided to them. Latino caregivers suggested additional avenues to spread awareness of the STAR-VTF program, such as working with community partners and using social media.

Sharing Information With Relatives

Multiple Latino caregivers shared that they valued the information provided in the training and that they shared this information with family members who assist in the care of the person living with dementia and with family members who are in a similar caregiving role. For example, one caregiver said:

> *So I shared the videos with a group chat for the seven-week period. It was a group chat of my caregivers for my Nana, including myself. So it's going to be my sisters, my two sisters, my two cousins, my father, my uncle, and my aunt. And all of us took part in that group chat, and I shared it with them every time I got a new video because I felt like it was something that we could all benefit from. We all share the same struggle right now. And even though we're not in the same place, as far as mentally, taking care of my grandmother, this is a step to getting us all on the same page. And it seemed like the right thing to do to share the information that I was getting and make sure that my family understood why I was doing what I was doing.* [ID 24]

Community Outreach Strategies

To further improve the STAR-VTF program outreach, some caregivers shared that it is important to partner with trusted Latino community organizations, community health workers, religious centers, and caregiving organizations. As one caregiver explains, to reach Latino communities, it is important to consider partners who have established trust in the community to establish the credibility of the program. The caregiver said:

> *So often Latinos they feel more trusted with people that they know, so family and friends or church or organizations. So if you can tap into those kind of communities to share it, I think, will probably be your best bet in getting the word out. Even my mom, she's very hesitant to go to the doctor. I have to take her because she grew up in a generation where they kind of mistrust doctors, but they trust their sisters or cousins. And if they come at them with information, they're more apt to use it than if it comes from a really formal setting. So to say that word of mouth, that community-centered family groups, community groups, churches, any kind of groups that have a good connection to the Latino community, I think they're more apt to use the information you're providing.* [ID 27]

Some caregivers also suggested advertising the intervention in medical facilities and partnering with health professionals. As one Latino caregiver shared:

> *Working with the community health workers, they can put it out (information on intervention) or maybe have something in the rural area at the doctors. I know they have a lot of doctors. So if it's just one doctor's office, if you have some handout or a little small TV screen that's on loop playing some of that information, it would be good.* [ID 09]

Some Latino caregivers also shared the importance of the STAR-VTF program building a digital presence by posting the training modules on social media, such as Facebook, to reach Latino populations. Caregivers believed delivering the STAR-VTF program through these platforms would increase the reach of the program. For example, one caregiver shared:

> *You can put them (training videos) on Facebook, on social media, and then when you're on social media, you could pop up and ... because you could be on social media on the weekends, and then your study would pop up, or your videos would pop up on the weekends and ... because people are always on social media day, morning, night, weekends, holidays, so if*

you had your content available on social media, then they would be able to view it. [ID 08]

Preliminary Outcomes

The average RMBPC score for a person living with dementia behavior problems changed from 28.24 (SD 18.31) at baseline to 21.34 (SD 16.76) postintervention, indicating a decreased frequency of behavior problems (Table 3). The average RMBPC score for the caregiver's reaction to behavior problems changed from 40.40 (SD 16.68) at baseline to 37.21 (SD 14.60) postintervention, indicating decreased caregiver reaction to behavior problems. Furthermore, the average Preparedness for Caregiving Scale score changed from 1.98 (SD 0.33) at baseline to 2.43 (SD 0.32) postintervention, indicating increased preparedness for caregiving. The findings show a trend toward decreased behavior problems in the person living with dementia, decreased caregiver reaction to behavior problems, and increased caregiver preparedness after versus before the intervention.

Table . Results for RMBPC[a] and Preparedness for Caregiving Scale at weeks 0 and 8.

	Week 0, mean (SD)	Week 8, mean (SD)
RMBPC		
Problem Frequency, overall (n=14)	28.24 (18.31)	21.34 (16.76)
Problem Frequency, memory subscale (n=14)	7.70 (5.35)	4.98 (3.63)
Problem Frequency, disruption subscale (n=14)	10.24 (7.95)	6.39 (6.20)
Problem Frequency, depression subscale (n=14)	10.73 (10.36)	9.23 (10.35)
Caregiver Reaction, overall (n=13)	40.40 (16.68)	37.21 (14.60)
Caregiver Reaction, memory subscale (n=13)	19.21 (5.70)	18.64 (6.52)
Caregiver Reaction, disruption subscale (n=13)	9.35 (6.54)	8.14 (5.27)
Caregiver Reaction, depression subscale (n=13)	11.85 (7.52)	10.43 (5.88)
Preparedness for Caregiving Scale Score (n=14)	1.98 (0.33)	2.43 (0.32)

[a]RMBPC: Revised Memory and Behavior Problem Checklist.

Discussion

Principal Results

This study aimed to pilot-test the culturally adapted STAR-VTF training modules among Latino caregivers of people living with dementia. Delivering the culturally adapted STAR-VTF training modules to Latino caregivers was feasible, with a high proportion of participants completing the study procedures. The training modules were also found to be easy for Latino caregivers to use based on the high average of System Usability Scale scores across modules.

The qualitative analysis of interviews with Latino caregivers who completed the 8-week program revealed their high satisfaction and acceptability with the training modules and components on dementia education, ABC problem solving, and caregiver support strategies. The study also identified areas for improvement in the training modules, such as providing caregivers with information on additional resources and enhancing the engagement of training videos. The study also provided valuable insights into expanding the intervention's reach, including through word of mouth, partnerships with community organizations, and social media. These findings will significantly contribute to the refinement of the culturally adapted STAR-VTF training modules for Latino caregivers. The culturally adapted STAR-VTF intervention was also found to have positive changes in the outcomes of behavioral problems among people living with dementia and caregiver preparedness among the small sample of Latino caregiver participants. Further research is needed to examine the efficacy of the culturally adapted STAR-VTF intervention.

Comparison With Prior Work

This is the first pilot study that tests a cultural adaptation of the STAR-VTF evidence-based intervention for Latino caregivers. Our findings on the feasibility of the culturally adapted STAR-VTF intervention are comparable to other interventions for dementia caregivers delivered digitally, which were also feasible [26,27]. However, unlike these studies, this study consists exclusively of Latino participants, which provides further information on the feasibility of digital interventions for Latino caregivers. Preliminary findings of positive mental health outcomes and improvements in caregiver preparedness are similar to findings from a systematic review of literature on internet-based interventions aimed at supporting family caregivers of people living with dementia [28]. The studies presented in the systematic review of digital dementia caregiving interventions reported reductions in caregiver stress, depression, and strain and improvements in self-efficacy [28]. However, conclusions from these studies and this pilot study should be made with caution due to study design limitations, as these studies consist of quasi-experimental, small randomized controlled trials, pre-post studies, and feasibility studies.

Our findings on acceptability and Latino caregivers' satisfaction with training components in this pilot study were consistent with the identified needs and wants of caregivers who participated in a qualitative study to test low-fidelity versions of the culturally adapted training modules [8]. In the pilot study, Latino caregivers expressed how the module on dementia education was especially beneficial to their understanding of dementia; this finding aligned with caregivers' perspectives of wanting more dementia education in the previous study of

low-fidelity training modules [8]. The desire to learn about ADRD from Latino caregivers of people living with dementia is also consistent with findings from a qualitative study that examined barriers and facilitators to increasing engagement of Hispanics or Latinos in clinical trials on ADRD [29]. In this study, a majority of caregivers expressed an eagerness to receive information about ADRD, including details on the different types of ADRD, specific risk factors for Latino communities, and hereditary characteristics [29].

Another aspect of the training modules that resonated with Latino caregivers, which was highlighted in our previous qualitative study, was the inclusion of reminders for caregivers to be empathetic toward people living with dementia. Our findings on caregiver satisfaction with the caregiver support strategy component are also similar to findings from our previous study in which caregivers highlighted the importance of self-care [8]. The need for information on self-care by Latino caregivers was also highlighted in a qualitative study that explored the needs of Latino caregivers of relatives with early-onset Alzheimer disease [30]. This study found that among their 27 participants, they overwhelmingly expressed the need for additional information concerning self-care, especially mental health care strategies and stress management [28]. Caregivers' high satisfaction with the training components piloted in this study suggests that we effectively integrated feedback from Latino caregivers to meet their needs in the training content.

Regarding our findings on areas of improvement, Latino caregivers' needs for additional dementia care–related resources are consistent with findings from a mixed methods study on the health information–seeking behavior of Latino caregivers of people living with dementia. The study on health-seeking behaviors found that Latino caregivers desire to seek various types of dementia information, including the availability of community resources, health care services, and treatment [31]. Caregivers in this pilot study also suggested improving training videos by making them more engaging with video reenactments of caregiving scenarios, knowledge checks, and caregiver testimonials. This aligns with findings from our prior qualitative study wherein Latino caregivers who viewed the low-fidelity modules expressed the idea of having real people in the videos to depict caregiver scenario examples [8]. Multiple studies have found that using video vignettes in the form of telenovelas and Spanish soap operas is an acceptable and effective method of delivering health information to Latino caregivers [32-34]. Furthermore, a study that evaluated the acceptability of a web-based course for caregivers of older adults found that caregivers liked having postmodule quizzes, and they also had a caregiver suggest the use of caregiver testimonials in course modules [35]. Our findings show that these preferences in web-based training are also warranted among a sample of Latino caregivers.

Limitations

Although preliminary outcomes were positive, this study had several limitations. First, due to the small budget, we were not able to pilot-test the full STAR-VTF intervention, which consists of weekly training modules and weekly 30-minute phone check-ins with a coach. It is during these check-ins that coaches review caregivers' ABC plans. Future pilot studies will include the coaching component, which will also ensure that the ABC plans are being developed appropriately. Second, general limitations of pilot studies apply, such as the inability to determine efficacy and estimate effect sizes for power calculations of a larger-scale study [17]. The target sample size of 20 for this pilot study was determined based on logistical constraints, as it was the maximum feasible within the available budget and study timeline. Third, the pre-poststudy design is subject to limitations due to the lack of a comparison group, which relies on conclusions formed from the temporal relationship of the measurements to the intervention. This type of study does not allow researchers to control for other changes that could have occurred around the same time as the intervention that could have influenced outcome results [36]. A limitation of the qualitative portion of this study is that caregivers had trouble recalling details about earlier training modules. A caregiver shared that since it had been over a month since they viewed the first couple of training modules, it was difficult to remember some of the specific details of the content covered. When conducting interviews with participants, the research assistants (CG and LZ) would sometimes need to remind caregivers about earlier training modules. The time gap between viewing the earlier training modules and being interviewed can result in recall bias, and participants may have omitted important details regarding their experience. Another limitation of this study is that 10 of the 14 interviews were conducted by the same research assistant who recorded the voiceover for the modules. If participants recognized that it was the same person, it may have led to social desirability bias, as participants may have provided answers that would be viewed favorably by the research assistant.

Conclusions

This pilot study found the culturally adapted STAR-VTF intervention to have high levels of feasibility, usability, and acceptability. In general, caregivers found the training modules beneficial and relatable and were highly satisfied with the cultural adaptations to the STAR-VTF intervention. The feedback from the qualitative interviews with Latino caregivers will help inform decisions during the fourth stage (Adaptation refinement) of the Barrera and Castro [16] framework on the cultural adaptation of evidence-based interventions to refine future iterations and meet Latino caregiver needs. Although future research is still necessary to test the efficacy of the culturally adapted STAR-VTF program, this pilot study provided valuable information about the feasibility, usability, and caregiver experiences with the intervention. This study also contributes to the broader knowledge of modifying evidence-based interventions to better serve Latino caregivers of persons with dementia.

Acknowledgments

This work was funded by the National Institute on Aging (NIA) of the National Institutes of Health (NIH) under award U54AG063546, which funds NIA Imbedded Pragmatic Alzheimer's and Alzheimer Disease–Related Dementias Clinical Trials Collaboratory (NIA IMPACT Collaboratory). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English and Spanish interview guides.
[DOCX File, 28 KB - aging_v8i1e66053_app1.docx]

References

1. Martinez IL, Gonzalez EA. Care v. caring: obligation, duty, and love among Latino Alzheimer's family caregivers. J Appl Gerontol 2022 Jul;41(7):1744-1751. [doi: 10.1177/07334648221084998] [Medline: 35484911]
2. Quiroz YT, Solis M, Aranda MP, et al. Addressing the disparities in dementia risk, early detection and care in Latino populations: highlights from the second Latinos & Alzheimer's symposium. Alzheimers Dement 2022 Sep;18(9):1677-1686. [doi: 10.1002/alz.12589] [Medline: 35199931]
3. Vega B, Vega WA, Wu S, Resendez J, Jin H, Aranda M. Latinos and Alzheimer's disease. Innov Aging 2018 Nov 1;2(suppl_1):216-216. [doi: 10.1093/geroni/igy023.794]
4. Balbim GM, Magallanes M, Marques IG, et al. Sources of caregiving burden in middle-aged and older Latino caregivers. J Geriatr Psychiatry Neurol 2020 Jul;33(4):185-194. [doi: 10.1177/0891988719874119] [Medline: 31510848]
5. Martinez IL, Gonzalez EA, Quintero C, Vania MJ. The experience of Alzheimer's disease family caregivers in a Latino community: expectations and incongruences in support services. J Gerontol B Psychol Sci Soc Sci 2022 Jun 1;77(6):1083-1093. [doi: 10.1093/geronb/gbab170] [Medline: 34543386]
6. Alzheimer's Association. 2023 Alzheimer's disease facts and figures. Alzheimers Dement 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]
7. Luchsinger JA, Tipiani D, Torres-Patiño G, et al. Characteristics and mental health of Hispanic dementia caregivers in New York City. Am J Alzheimers Dis Other Demen 2015 Sep;30(6):584-590. [doi: 10.1177/1533317514568340] [Medline: 25635108]
8. Garcia CN, Duran MC, Ramirez M. Refining cultural adaptations of a behavioral intervention for Latino caregivers of people living with dementia: qualitative interview study in Washington State. JMIR Aging 2024 Jan 11;7(1):e53671. [doi: 10.2196/53671] [Medline: 38206663]
9. McCurry SM, Logsdon RG, Mead J, et al. Adopting evidence-based caregiver training programs in the real world: outcomes and lessons learned from the STAR-C Oregon translation study. J Appl Gerontol 2017 May;36(5):519-536. [doi: 10.1177/0733464815581483] [Medline: 25873454]
10. Lindauer A, Croff R, Mincks K, et al. "It took the stress out of getting help": the STAR-C-telemedicine mixed methods pilot. Care Wkly 2018;2:7-14. [doi: 10.14283/cw.2018.4] [Medline: 30393783]
11. Teri L, McCurry SM, Logsdon R, Gibbons LE. Training community consultants to help family members improve dementia care: a randomized controlled trial. Gerontologist 2005 Dec;45(6):802-811. [doi: 10.1093/geront/45.6.802] [Medline: 16326662]
12. Teri L, McKenzie G, Logsdon RG, et al. Translation of two evidence-based programs for training families to improve care of persons with dementia. Gerontologist 2012 Aug;52(4):452-459. [doi: 10.1093/geront/gnr132] [Medline: 22247431]
13. Lindauer A, McKenzie G, LaFazia D, et al. Using technology to facilitate fidelity assessments: the Tele-STAR caregiver intervention. J Med Internet Res 2019 May 24;21(5):e13599. [doi: 10.2196/13599] [Medline: 31127721]
14. Satz S, White Q, Watkin LG. Utilizing technology and integrated services to support family caregivers in rural communities. Innov Aging 2023 Dec 21;7(Supplement_1):618-618. [doi: 10.1093/geroni/igad104.2015]
15. Ramirez M, Duran MC, Penfold RB, Pabiniak CJ, Hansen KE, Ornelas IJ. STAR-Caregivers Virtual Training and Follow-up: a cultural adaptation for Latino caregivers of people with dementia. Transl Behav Med 2023 Sep 12;13(9):625-634. [doi: 10.1093/tbm/ibad028] [Medline: 37130336]
16. Barrera M, Castro FG. A heuristic framework for the cultural adaptation of interventions. Clin Psychol Sci Pract 2006;13(4):311-316. [doi: 10.1111/j.1468-2850.2006.00043.x]
17. Pilot studies: common uses and misuses. NCCIH. URL: https://www.nccih.nih.gov/grants/pilot-studies-common-uses-and-misuses [accessed 2024-06-05]
18. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation In Industry: CRC Press; 2014:4-7.

19. Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist. Psychol Aging 1992 Dec;7(4):622-631. [doi: 10.1037//0882-7974.7.4.622] [Medline: 1466831]
20. Salvia MG, Dawidowski A, Schapira M, et al. Spanish Revised Memory and Behavior Problems Checklist Scale (SpRMBPC): trans-cultural adaptation and validation of the RMBPC questionnaire. Int Psychogeriatr 2011 Sep;23(7):1160-1166. [doi: 10.1017/S1041610210002073] [Medline: 21294937]
21. Preparedness for Caregiving Scale. HIGN. URL: https://hign.org/consultgeri/try-this-series/preparedness-caregiving-scale [accessed 2024-06-05]
22. Gutierrez-Baena B, Romero-Grimaldi C. Development and psychometric testing of the Spanish version of the Caregiver Preparedness Scale. Nurs Open 2021 May;8(3):1183-1193. [doi: 10.1002/nop2.732] [Medline: 33340272]
23. Sevilla-Gonzalez MDR, Loaeza LM, Lazaro-Carrera LS, et al. Spanish version of the System Usability Scale for the assessment of electronic tools: development and validation. JMIR Hum Factors 2020 Dec 16;7(4):e21161. [doi: 10.2196/21161] [Medline: 33325828]
24. Miles MB, Huberman AM, Saldana J. Qualitative Data Analysis: A Methods Sourcebook: Sage Publishing; 2014.
25. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J User Exper 2009;4(3):114-123 [FREE Full text]
26. Loi SM, Tropea J, Gaffy E, et al. START-online: acceptability and feasibility of an online intervention for carers of people living with dementia. Pilot Feasibility Stud 2022 Feb 16;8(1):41. [doi: 10.1186/s40814-022-00999-0] [Medline: 35172894]
27. Kwok I, Lattie EG, Yang D, et al. Acceptability and feasibility of a socially enhanced, self-guided, positive emotion regulation intervention for caregivers of individuals with dementia: pilot intervention study. JMIR Aging 2023 Sep 6;6(1):e46269. [doi: 10.2196/46269] [Medline: 37672311]
28. Hopwood J, Walker N, McDonagh L, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216. [doi: 10.2196/jmir.9548] [Medline: 29895512]
29. Marquez DX, Perez A, Johnson JK, et al. Increasing engagement of Hispanics/Latinos in clinical trials on Alzheimer's disease and related dementias. Alzheimers Dement (N Y) 2022;8(1):e12331. [doi: 10.1002/trc2.12331] [Medline: 35910673]
30. Withers M, Cortez-Sanchez K, Herrera J, Ringman JM, Segal-Gidan F. "My backpack is so heavy": experiences of Latino caregivers of family with early-onset Alzheimer's. J Am Geriatr Soc 2021 Jun;69(6):1539-1547. [doi: 10.1111/jgs.17091] [Medline: 33713418]
31. Ramirez M, De Anda S, Jin H, Herrera JR, Wu S. Health information-seeking behavior of Latino caregivers of people living with dementia: a mixed-methods study. J Appl Gerontol 2023 Aug;42(8):1738-1748. [doi: 10.1177/07334648231163430] [Medline: 36932723]
32. Cruz-Oliver DM, Malmstrom TK, Fernández N, Parikh M, García J, Sanchez-Reilly S. Education intervention "Caregivers Like Me" for Latino family caregivers improved attitudes toward professional assistance at end-of-life care. Am J Hosp Palliat Care 2016 Jul;33(6):527-536. [doi: 10.1177/1049909115584315] [Medline: 26019262]
33. Gallagher-Thompson D, Tzuang M, Hinton L, et al. Effectiveness of a fotonovela for reducing depression and stress in Latino dementia family caregivers. Alzheimer Dis Assoc Disord 2015;29(2):146-153. [doi: 10.1097/WAD.0000000000000077] [Medline: 25590939]
34. Kajiyama B, Fernandez G, Carter EA, Humber MB, Thompson LW. Helping Hispanic dementia caregivers cope with stress using technology-based resources. Clin Gerontol 2018;41(3):209-216. [doi: 10.1080/07317115.2017.1377797] [Medline: 29236621]
35. Rottenberg S, Williams A. Web-based delivery of the caregiving essentials course for informal caregivers of older adults in Ontario: mixed methods evaluation study. JMIR Aging 2021 Jun 15;4(2):e25671. [doi: 10.2196/25671] [Medline: 34128815]
36. Aggarwal R, Ranganathan P. Study designs: part 4—interventional studies. Perspect Clin Res 2019;10(3):137-139. [doi: 10.4103/picr.PICR_91_19] [Medline: 31404185]

Abbreviations

ABC: activator-behavior-consequence
ADRD: Alzheimer and related dementia
REDCap: Research Electronic Data Capture
RMBPC: Revised Memory Behavior Problem Checklist
STAR-C: STAR Caregivers
STAR-VTF: STAR Virtual Training Follow Up

Edited by A Quialheiro; submitted 04.09.24; peer-reviewed by A Lindauer, K Meyer; revised version received 03.02.25; accepted 21.04.25; published 10.06.25.

Please cite as:
Mariscal MA, Garcia C, Zavala L, Ramirez M
Culturally Adapted STAR-Caregivers Virtual Training and Follow-Up for Latino Caregivers of People Living With Dementia: Single-Arm Pre-Post Mixed Methods Study
JMIR Aging 2025;8:e66053
URL: https://aging.jmir.org/2025/1/e66053
doi:10.2196/66053

Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool

Olga A Biernetzky[1], BA, MSc, PhD; Jochen René Thyrian[2,3], Prof Dr rer med; Melanie Boekholt[2]; Matthias Berndt[2]; Wolfgang Hoffmann[2,3], MPH, Prof Dr med; Stefan J Teipel[1,4], Prof Dr med; Ingo Kilimann[1,4], Dr med

[1]Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Gehlsheimer Strasse 20, Rostock, Germany

[2]Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Greifswald, Germany

[3]Institute for Community Medicine, Section Epidemiology and Community Health, University Medicine Greifswald, Greifswald, Germany

[4]Department for Psychosomatic and Psychotherapeutical Medicine, Rostock University Medical Center, Rostock, Germany

Corresponding Author:
Olga A Biernetzky, BA, MSc, PhD
Deutsches Zentrum fuer Neurodegenerative Erkrankungen, Rostock/Greifswald, Gehlsheimer Strasse 20, Rostock, Germany

Abstract

Background: Despite the increasing interventions to support family caregivers of people with dementia, service planning and delivery is still not effective.

Objective: Our study aimed to develop a digitally-supported needs assessment tool for family caregivers of people with dementia that is feasible, time-efficient, understood by users, and can be self-completed in the primary care setting.

Methods: The development of the unmet needs assessment tool was part of a cluster-randomized controlled trial examining the effectiveness of a digitally supported care management programme to reduce unmet needs of family caregivers of people with dementia (GAIN [Gesund Angehörige Pflegen]) and was conducted in 3 phases. Using an iterative participatory approach with informal caregivers, health care professionals including general practitioners, neurologists, psychologists, psychiatrists, nurses, and Alzheimer Society representatives, we developed a digital self-completion unmet needs assessment tool focusing on informal caregivers' biopsychosocial health und quality of life in connection to their caregiver responsibilities. Data were collected through group discussions, written feedback, protocols, think-aloud protocols, and interviews, and analyzed thematically.

Results: Data from 27 caregivers, including caregivers of people with dementia (n=18), health care professionals (n=7), and Alzheimer Society representatives (n=2) were collected. Thematic analysis identified 2 main themes: content of the assessment tool and usability and handling of the digital tablet-based assessment tool. The feedback provided by the stakeholders led to new aspects and changes to make the tool comprehensive, easy to read, and easy to handle. The overall mean completion time was reduced from the initial 37 minutes to 18 minutes, which renders the assessment tool fit to be self-completed in waiting rooms of primary care practices or other settings.

Conclusions: The input of the 3 stakeholder groups has supported the development of the assessment tool ensuring that all aspects considered important were covered and understood and the completion of the assessment procedure was time-efficient and practically feasible. Further validation of the assessment tool will be performed with the data generated as part of the GAIN trial.

Trial Registration: ClinicalTrials.gov NCT04037501; https://clinicaltrials.gov/study/NCT04037501

(JMIR Aging 2025;8:e59942) doi:10.2196/59942

KEYWORDS

unmet needs; assessment development; family caregivers of people with dementia; dementia; need; Alzheimer; self-guided; self-reported; caregiver; informal care; spousal care; interview; qualitative; thematic; usability; mHealth; tablet; self-completed; aging; patient care; health interventions; care giver; digital health; ehealth; digital assessment; memory

Introduction

The World Health Organization (WHO) estimates 55 million people are living with dementia worldwide, with an increase to 78 million by 2030 [1]. In Germany, approximately two-thirds of the 1.8 million people with dementia are cared for at home by an informal caregiver, most often a relative [2]. Numerous studies have shown that caring for a relative with dementia is associated with a multitude of time- and resource-intensive challenges [3-5]. Emotional stress, social isolation, depression,

and financial burden are among the negative consequences that dementia family caregivers have reported [6-10]. To alleviate these consequences, interventions for the reduction of caregivers' burden have been developed [11-14], such as telephone-based as well as multicomponent psychosocial and individualized support programs [11,13]. With the growing consensus that caregiver support is an integral part of the care of patients with dementia, there was also an increase in interventions to support informal caregivers [12]. However, each caregiver's situation is different and requires an individual set of interventions. Individualized interventions are still underused in care practice as they require time and dementia-specific expertise that is often lacking in primary care. A systematic and comprehensive yet individualised unmet needs assessment tool would be useful for health care professionals (HCPs) to better plan services and interventions in clinical settings as it would flag areas where the individual caregiver needs support.

In a mixed methods study, Stirling et al [15] used Bradshaw taxonomy of need [16] to explore the relationship between different types of caregiver service need. Bradshaw and Care [16] defined categories of need as measures of professionally identified carer burden (normative need), service use (expressed need), carer's stated need (felt need) and the comparison of groups using services with groups who do not (comparative need). It is argued that no single measure of the above needs is likely to capture all carer's unmet needs. In the light of person-centred care, the normative need has been increasingly challenged as paternalistic and inappropriate [15,17].

Therefore, a focus on caregivers' stated need (felt need) as well as service use (expressed need) would be an appropriate starting point to consider for an unmet needs assessment tool. An individualized, digitally supported self-assessment instrument can aid in the efficient identification of health and support areas that should be addressed from the perspective of the informal caregivers. This allows collaborative intervention planning and reduces the risk of a paternalistic doctor-patient relationship [18]. If a digital system is to truly ensure collaborative service planning, the digital assessments instrument must be developed with the participation of the users themselves.

The direct input of family caregivers and other stakeholders into the content, structure, and handling of the digital assessment system provides crucial feedback regarding its content validity as well as comprehensiveness and technical usability. Through a greater level of involvement of stakeholders, health care services can be improved and be more applicable in real-world clinical settings and facilitate implementation [19]. A systematic review of instruments used to assess the needs of family dementia caregivers identified 36 instruments that were described in detail or in part in the 70 publications included [20]. The authors reported that only one instrument was partially validated to assess the needs of family dementia caregivers, namely, the Carer's Needs Assessment for Dementia (CNA-D), a semistructured interview not intended for clinical use with a completion time of about an hour. For conceptual clarity, researchers recommended establishing a theoretical model or framework to organize the diverse needs of family dementia caregivers [21]. In addition, researchers argue that a focus on the detection of changes in family caregivers' needs throughout disease progression is important [20,22].

Therefore, the aim of this study was to develop a digital unmet needs assessment tool that can be used in primary care settings. We conducted the study in collaboration with informal and formal caregivers and other stakeholders. Our objectives were to (1) create a user-friendly and family dementia caregiver specific digitally supported assessment tool, (2) develop an assessment process that is understandable and comprehensive yet time-efficient, and (3) create an assessment tool that can be used at multiple locations such as general practitioners' (GPs) offices and memory clinic waiting rooms.

Methods

Overview

The development of the assessment tool was in preparation of the upcoming randomized controlled intervention trial examining the effectiveness of a digitally supported care management intervention for family caregivers (German title: Gesund Angehörige Pflegen [GAIN]). The trial was supported by the Federal Joint Committee (G-BA [Gemeinsamer Bundesausschuss/Innovationsausschuss]). The funding code (FKZ) is 01VSF18030).

Participants

We involved individuals from 3 stakeholder groups, namely, family dementia caregivers, health care professionals (general practitioners, neurologists, psychiatrists, psychologists, and nurses), and Alzheimer Society representatives. We approached HCPs of a medical university and memory clinic staff, as well as Alzheimer Society representatives, and asked for their feedback (phase 1). Family dementia caregivers who were visiting a memory clinic with their relative or friend with dementia were asked whether they would like to provide feedback on the paper-based version of the assessment instrument regarding the content, clarity, and comprehensiveness (phase 2) as well as the usability of the digital version of the assessment tool as a self-completion assessment instrument on a PC tablet (phase 3). No further demographic data were collected.

Procedure

We used an iterative user-centered participatory approach with 3 phases. Previously used questionnaires [20,23,24] assessing unmet needs and care interventions from previous dementia care management studies [4,25-31] in Germany were used as starting point.

Phase 1

In phase 1, the authors developed a list of previously used questionnaires assessing unmet needs in family dementia caregivers as well as a list of available care interventions offered in Germany. The selection of questionnaires was based on a comprehensive review of the literature covering the databases OVID, MEDLINE, and PsycInfo searching for unmet needs, caregiver needs assessments methods and family dementia caregivers. With respect to available care interventions, we used work from previous dementia care management studies

management studies [4,25-31] offering a list of interventions to family dementia caregivers in Germany. We used Germany-based intervention studies as these are health care system specific. Criteria for the inclusion of questionnaires were the content they covered (eg, quality of life, psychosocial factors, and health-related domains), their previous use in research and interventions, and their completion time. We also considered questionnaires based on either their recommendation by the EU Joint Programme in Neurodegenerative Research (JPND) Working Group on Longitudinal Cohorts or common use in larger German trials such as IDemUck [32], DelpHi-MV (Dementia: life- and person-centered help in Mecklenburg-Western Pomerania) [26,28], intersec-CM [33,34], or DemNet-D [35,36]. These questionnaires are validated and allowed comparison with German and international studies. The authors then further discussed these selections in meetings, compared 5 separate versions of a possible unmet needs assessment instrument and checked whether items would cover interventions offered as part of the dementia care management conducted in previous primary care studies [4,26,28,37,38]. With this approach, we wanted to ensure that the operationalization of problem-centered needs were connected to available services within the German health care system and could therefore be addressed in this framework accordingly.

The authors selected different assessment versions aiming to cover all biopsychosocial aspects and compared different scales for caregiver burden. The need items were then compared and matched with the list of informal caregiver interventions from previous studies—these were marked according to a match or no match. This allowed the calculation of percentages of “items covered” from the dementia caregiver intervention list for each version of questionnaires. The five versions of combinations of questionnaires were checked by 3 authors and researchers for plausibility. The results of this item-intervention matching process were used to identify which versions covered most unmet needs with matched possible interventions. These are shown in Table 1.

Table . Proportion of match with family dementia caregiver interventions.

Versions and interventions		Match, %
Version 1		47
	FIMA[a]	
	URN[b]	
	EQ-5D-5L[c]	
	ZBI-7[d]	
	LSNS-6[e]	
Version 2		17
	FIMA	
	URN	
	EQ-5D-5L	
	BIZA-D[f]	
Version 3		96
	CANE[g]	
	EQ-5D-5L	
	ZBI-7	
	LSNS-6	
Version 4		94
	CANE	
	DQoL-OC[h]	
Version 5		39
	HABC[i]	
	DQoL-OC	
	ZBI-7	

[a]FIMA: Questionnaire for the Use of Medical and Non-Medical Services in Old Age [39].

[b]URN: Caregiver unmet resource needs scale [40].

[c]EQ-5D-5L: Health-related quality of life [41].

[d]ZBI-7: Zarit Burden Interview [42,43].

[e]LSNS-6: Lubben Social Network Scale [44,45].

[f]BIZA-D: The Berlin inventory of the burden on relatives - dementia – Module 3,5,and 6 [46].

[g]CANE: Camberwell Assessment of Need for the Elderly [47,48].

[h]DQoL-OC: The Dementia Quality of Life Scale for Older Family Carers [49].

[i]HABC: Healthy Aging Brain Care Monitor—Caregiver Version [50].

Versions 3 and 4 covered most items, whereas versions 1, 2, and 5 covered less than 50% of the dementia-specific items. The two versions covering most unmet needs were then presented and discussed at an advisory board meeting with input from HCPs (eg, neurologists, psychiatrists, psychologists, and nurses) and Alzheimer Society representatives. Meeting minutes were recorded of all points raised and discussed. This input was used to select the final version of the assessment tool.

Phase 2

In phase 2, the chosen assessment version covering most unmet needs was tested in a memory clinic. A tablet computer was deemed most appropriate to serve as digital device due to its size, weight, flexibility, and ease of handling. Based on a previous study on the use of a tablet-based digital expert system [26,28,30], we had taken several decisions on the hardware basis of our system before the participatory part of our study started. The content and usability aspects, however, were developed together with future users. The authors approached family dementia caregivers in a memory clinic to provide feedback and suggestions on a paper-based version of the assessment tool. The changes were primarily focused on the Camberwell Assessment of Need for the Elderly (CANE) questionnaire, as well as the demographic and informal caregiver-specific questions, as these were items that could still be adapted. Family caregivers who were interested in

contributing were given a short version comprising 7 questions of the CANE questionnaire. The researcher noted the caregivers' comments and thoughts in a think-aloud protocol. Together with a researcher, they were asked to read the questions and share their thoughts with respect to each question and their understanding thereof. The focus was on comprehension, words and sentence structure, answer options, descriptions, and additional comments. In the second round, both informal caregivers and health care professionals were approached to provide feedback on the whole printed version of the assessment instrument. Those who agreed to provide feedback were asked to mock-complete the assessment tool and write down their notes and questions. In this round, informal caregivers and HCPs were asked to measure the time needed to complete the assessment procedure. We provided written questions such as "Are the questions clear and easy to understand?"; "Where did you encounter problems when completing the assessment?"; "Did you have issues understanding the questions?"; and "Are there questions or areas that should be in the assessment set but are currently missing?" Based on the feedback received, the first draft of the assessment set was modified resulting in a second draft. When adapting the first draft, our focus was on the assessment procedure being clear and easy to understand for family caregivers in terms of the way questions are phrased. Based on the feedback we received, several parts of the assessment set were adapted and duplicates were removed by the authors resulting in the final draft.

Phase 3

In phase 3, a digitalized tablet-based version of the final version of the assessment tool was tested for its technical usability. In this phase, we asked both informal caregivers and HCPs to test the digital assessment tool with a focus on the handling of the tablet. Data was collected through interviews and think-aloud protocols. Family caregivers in a memory clinic were asked to complete the digital assessment tool in the waiting room and share their thoughts and comments with a study nurse or researcher who protocolled the comments. Similarly, HCPs of a medical university and a memory clinic were asked to complete the digital assessment procedure and write down feedback on the handling and any possible bugs or problems they noticed while testing the digital system.

Analysis

All data collected through written feedback, protocols, think-aloud protocols, and interviews were analyzed thematically following the steps by Braun and Clarke [51]. The six phases involved were (1) familiarizing ourselves with the data, (2) generating initial codes, (3) constructing themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report [51].

Ethical Considerations

The study is conducted in accordance with the criteria (valid at present) of the Declaration of Helsinki, the ICH-guidelines for Good Clinical Practice, the Memorandum for Safeguarding Good Scientific Practice (German Research Foundation), and the International Ethical Guidelines for Biomedical Research Involving Human Subjects. Ethical approval has been obtained from the Ethical Committee of the University Medicine Greifswald (BB120/2019) and the Ethical Committee of the University Medicine Rostock (A2020/0013). The trial is registered at ClinicalTrials.gov (NCT04037501).

Results

Results of the development of the assessment tool of the 3 phases are divided in the subsections presented below. The process of phases 2 and 3 was iterative comprising several rounds through which we collected feedback, made changes based on feedback received and then moved to the next round of collecting feedback.

Participants

A total of 27 family caregivers of people with dementia (n=18), HCPs (n=7), and Alzheimer Society representatives (n=2) participated. Among the HCPs were general practitioners, neurologists, psychologists, psychiatrists, and nurses. The mean age of the 27 family caregivers was 57 (SD 13.78; range 35‐88) years and 9 (50%) were women. The HCPs' and representatives' mean age was 41 (SD 8.05; range 28‐50) and 6 out of 9 (67%) were women. As this was a participatory approach, no further demographic data were collected.

Phase 1: Structure and Content of the Assessment Tool

After the input from the advisory board including neurologists, psychiatrists, psychologists, nurses, and Alzheimer Society representatives, the final structure of the assessment instrument was selected. It was the one with the highest percentage of matches between problem-centered needs and the associated interventions (69/72, 96%; version 3; see Table 1).

The unmet needs assessment tool comprised several validated questionnaires with a total of 57 questions examining a person's self-reported demographic information, use of medical and nonmedical services, unmet needs (CANE), health-related quality of life (EQ-5D-5L), caregiver burden (Zarit Burden Interview [ZBI]), and social support (Lubben Social Network Scale [LSNS]). It can be completed using a digital version or a paper-based version. In total, the assessment set covers 51 health- and support-related aspects.

Domains covered by the unmet needs assessment tool are the following: (1) health and care, (2) employment, (3) information and knowledge, (4) emotional support, (5) social support, and (6) caregiver burden.

The structure of the unmet needs assessment tool can be seen in Table 2.

Table . Structure of the unmet needs assessment tool.

Instrument	Domain
Socio-demographic information	—[a]
Caregiver specific information	Individual situation of informal caregiver
CANE[b]	Unmet needs I—Information, physical, mental
EQ-5D-5L	Health-related quality of life
LSNS[c]	Unmet needs II—Social network
ZBI[d]	Unmet needs III—Caregiver burden

[a]Not available.

[b]CANE: Camberwell Assessment of Need for the Elderly.

[c]LSNS: Lubben Social Network Scale.

[d]ZBI: Zarit Burden Interview.

The number of unmet needs addressed the participants' medical needs, home care needs, psychosocial needs, and needs connected to the caregiver role. This needs assessment instrument included selected parts of the CANE [47].

Health-related quality of life was assessed using the EQ-5D-5L [41]. This instrument comprises 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension has 5 levels varying from no problems to extreme problems. Each level corresponds to a 1-digit number that expresses the level selected for that dimension ranging from 1 to 5 with higher numbers indicating more severe problems. The digits for the 5 dimensions can be converted into an overall index by an not publicly accessible algorithm. The index describes the participant's health-related quality of life. It also contains a vertical visual analogue scale (EQ VAS), with endpoints that are labelled "Best imaginable health state" and "Worst imaginable health state." The VAS reflects the patient's own judgement and can be used as a quantitative measure of health outcome.

Social support was assessed using LSNS-6 [44,45]. This scale is a self-report measure of social engagement including contact to and interaction with family and friends on a 6-item scale. Total scores range from 0 to 30 with an equally weighted sum of the 6 items. The family and friends subscales include questions regarding the number of friends and family one has regular contact with as well as their availability for help and support in private matters. High scores indicate strong social networks.

Informal caregiver burden was assessed using the 7-item version of the Zarit-Burden Interview (ZBI-7). The short version ZBI is a caregiver self-report measure to examine burden, which is associated with functional and behavioral impairments in the social, psychological and physiological context, and home care situation [42,43]. It contains 7 items using a 5-point scale. Response options range from 0 (Never) to 4 (Nearly to Always). Total scores range from 0 indicating no burden to 28 indicating severe burden.

Phases 2 and 3: Feedback Received on the Assessment Tool

Themes

A total of two main themes were derived from the analysis: (1) content of the assessment instrument and (2) digital tablet usability for the assessment tool. Each theme had further subthemes and subtheme categories and were divided in either informal caregivers or HCPs and Alzheimer Society representatives as these groups had different areas they paid special attention to. Almost all subthemes had 1 to 3 subtheme categories, see Table 3.

Table . Themes, subthemes, and subtheme categories.

Theme and subtheme	Subtheme category
1. Content of the assessment procedure	—[a]
Informal caregivers	—
1.1 Comprehension	• 1.1.1 Simplification of questions • 1.1.2 Simplification of answer options • 1.1.3 Separate questions for caregiver health and caregiver care responsibilities in relation to the person with dementia
1.2 Assessment structure	• 1.2.1 No multiple questions in a table • 1.2.2 Questions whether help is needed in a specific domain for clear identification of unmet needs
1.3 Reduction of completion time	• 1.3.1 Fatigue
1.4 Areas important to caregivers	—
Health care professionals	
1.1 Assessment structure	• 1.1.1 Order of instruments • 1.1.2 Focus on caregivers' health and care responsibilities
1.2 Reduction of completion time	—
2. Digital tablet usability of the assessment tool	—
Informal caregivers	
2.1 Simple layout	• 2.1.1 Large font and buttons
2.2 Handling	• 2.2.1 Manual answer selection
Health care professionals	
2.1 Simple layout	• 2.1.1 One question per screen • 2.1.2 Option to go back and forth within the assessment procedure • 2.1.3 Progression bar
2.2 Handling	—

[a]Not applicable.

Content of the Assessment Instrument

Regarding the assessment's content, informal caregivers provided information on 4 subthemes regarding the comprehension, assessment structure, reduction of completion time, and areas important to them.

Comprehension (Informal Caregiver Perspective)

Regarding the clarity of the assessment procedure to them, caregivers commented on the simplification of questions and answer options. Instead of "This is an unmet need" or "This is a need met" (as suggested in the CANE) they preferred to answer "Yes" or "No," for instance. They also commented on being sometimes confused whether a question was referring to themselves or to the person diagnosed with dementia indicating that questions had to be framed specifically.

Assessment Structure

Regarding the assessment's structure, informal caregivers commented that multiple questions in a table format were confusing. In addition, some commented that they would like to self-identify the areas where they needed help rather than having professionals assume that they needed help or support in an area, which was incorporated into the assessment procedure by first asking whether there is a problem in a specific area and secondly asking whether the respondent receives enough help with respect to this problem. A yes or no response to these questions would then indicate the presence or absence of an unmet need.

HCPs and Alzheimer Society representatives considered the order of instruments within the assessment procedure and argued that the instrument covering most unmet needs (CANE) should be moved to the front to prevent that important health care domains received less attention due to fatigue. They also reported that the informal caregivers' own health as well as factors influencing their health due to their caregiver role should be a focus since a caregiver's unmet needs can arise from two sources: either from the personal needs of the family caregiver or from his or her care responsibilities for the people with dementia.

Reduction of Completion Time (Informal Caregiver Perspective)

The majority of the informal caregivers commented on the length of the initial versions of the assessment tool and reported

that they became a bit tired towards the end of the assessment procedure. A caregiver stated: "For the last few questions my focus was a bit lost" (informal caregiver, woman).

Areas Important to Caregivers (Informal Caregiver Perspective)

Caregivers commented on some of the areas that they found particularly important such as psychological support, mobility, social life, and psychoeducation.

A caregiver stated:

> *I would say that there should be a question for psychological support. Self-help groups should be offered locally and it should be asked whether advice or support is needed in this regard. Another thing that I would still find important is to ask how mobile you are and whether you need help in this regard. For example, I don't have a car and have to use public transport to get everywhere. And there are also people who don't even have a driver's license. And if you have to go somewhere and have to pick up things from the pharmacy (eg, incontinence pads), these are sometimes large packages that are also heavy and you have to get them home first. It would be a good idea to ask whether you need help in this regard.* [Informal caregiver, woman]

Another caregiver reported:

> *The social life is very important, my father is hardly involved in it anymore. Because of his swallowing problems, he can no longer even go to a restaurant, which is difficult.* [Informal caregiver, man]

Another cargiver stated:

> *Since we have not yet come into contact with services, I would like a lot of information so that you know what you have to do. And also so that you know why behavior changes, how people then think and what is going on inside you. I'd be very interested in that.* [Informal caregiver, man]

Regarding the content of the assessment instrument, HCPs and Alzheimer Society representatives provided information on 2 subthemes, namely, assessment structure and reduction of completion time.

Digital Tablet Usability of the Assessment Tool

Regarding the assessment's tablet usability, informal caregivers, HCPs, and Alzheimer Society representatives identified 2 subthemes, namely, simple layout and handling with slightly different subtheme categories.

Simple Layout

Informal caregivers commented on the font size, often taking out their reading glasses to be able to read the questions.

HCPs commented that one question per screen would allow the large font and button size needed. Some clinicians stated that it would be helpful to allow caregivers to go back within the assessment procedure in case they changed their mind regarding a previous answer. Going forward should only be possible if the question on the screen was completed. The majority of HCPs stated that an indication bar within the assessment procedure would be helpful for caregivers to receive information on their progress as well as how much of the assessment procedure is still in front of them.

Handling

The majority of informal caregivers reported that selecting answers by clicking worked well after adjustments were made to the font size and the size of the answer option buttons. Regarding the handling, all HCPs and Alzheimer Society representatives agreed that the selection of answers should be as straight forward as possible allowing caregivers to complete the assessment procedure without technical issues.

The adaptations made to the digital tablet-based assessment tool are shown in Figures 1 and 2. Figure 1 shows the initial, first version of the tablet-based assessment tool. Figure 2 shows the final tablet-based assessment version after the feedback was incorporated into the development by the IT team.

Figure 1. Outline of the first digital assessment version.

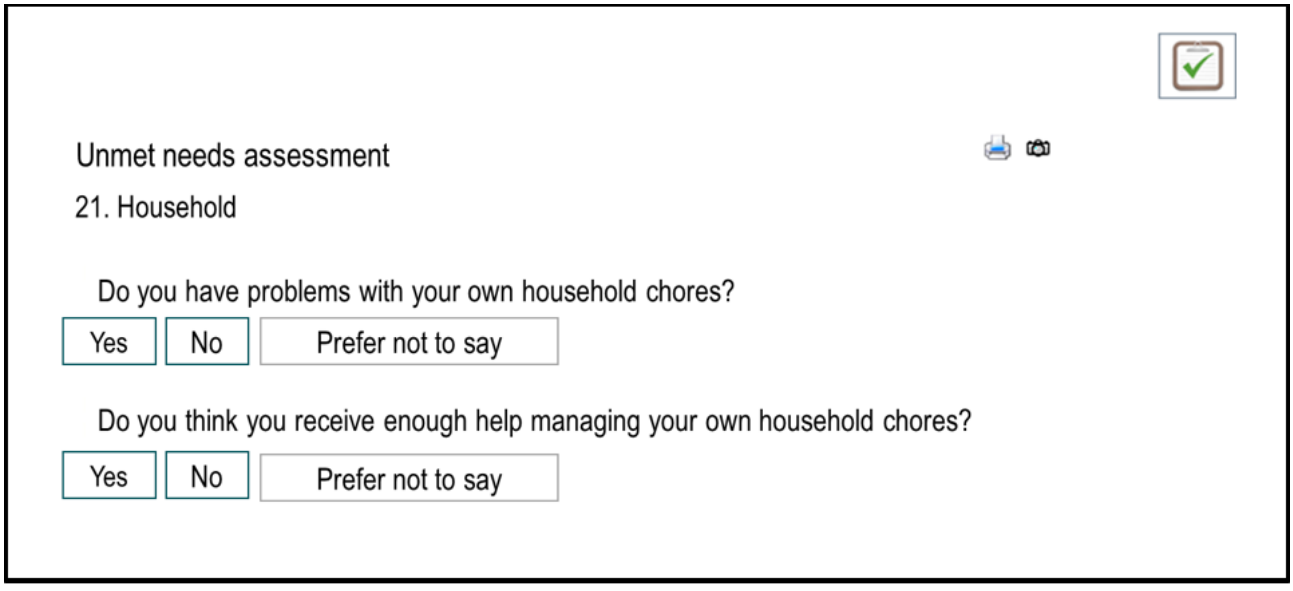
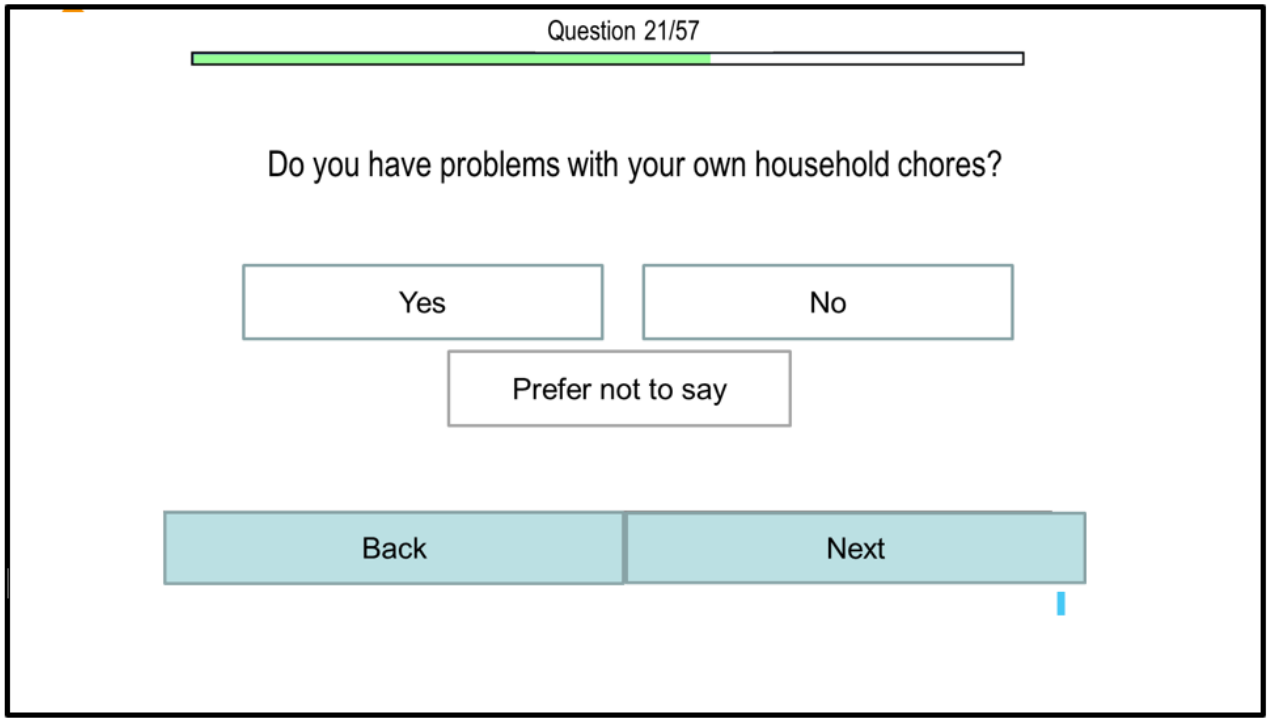


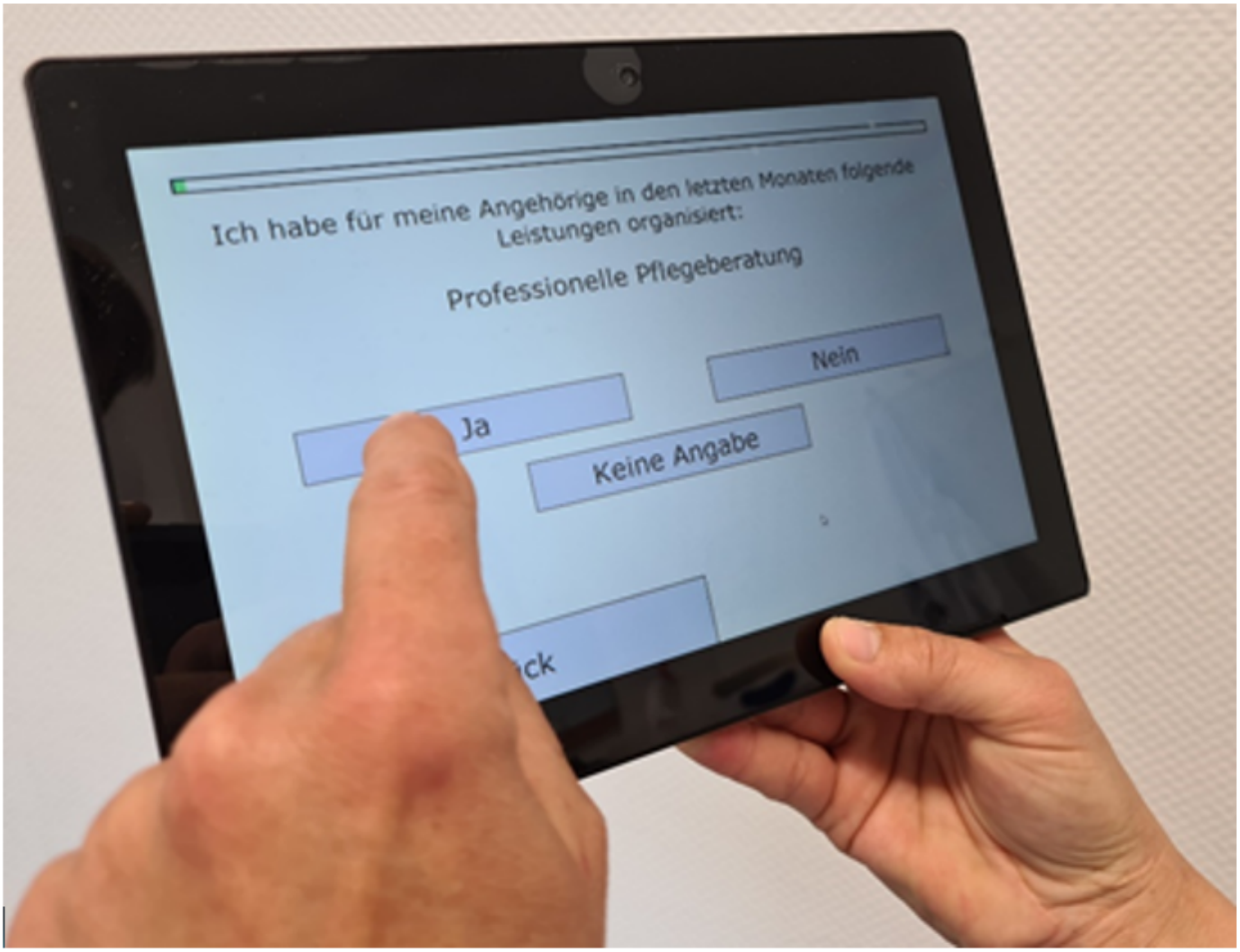
Figure 2. Outline of the final digital assessment version.



After completion, the outcome of the assessment procedure can be viewed by the study nurse as well as the GP. These professionals can then use the outcome to work together with the informal caregiver toward a reduction of the unmet needs identified. The study nurses provide the informal caregivers with a summary of the consultation where points discussed are listed.

Figure 3 depicts a caregiver filling out the digital version of the assessment tool on a tablet.

Figure 3. An informal caregiver completing the digital tablet-based assessment.



Discussion

Principal Findings

This study describes the phases of an iterative, participatory process we used to develop a digital tablet-based assessment tool designed to comprehensively identify the needs of informal caregivers of people with dementia in clinical practice. Through the collaboration with family dementia caregivers, health care professionals, and Alzheimer Society representatives, we developed a digital unmet needs assessment instrument that consists of a set of standardized instruments but also includes additional questions and adaptations that were incorporated along the iterative development process. The unmet needs assessment tool focuses on needs that can arise from two sources: either from the personal needs of the family caregiver or from the care responsibilities for people with dementia. With the contribution of informal caregivers and other stakeholders, we designed an assessment tool that covers a comprehensive range of needs considered important. At the same time, with the input of the stakeholders, the time to complete the assessment procedure was substantially reduced from the initial 37 minutes to 18 minutes. This indicates that we reached our goal of developing a comprehensive, feasible, and time-efficient assessment device.

With respect to the content of the assessment instrument, we identified important aspects regarding clarity, assessment structure, the reduction of completion time as well as content areas important to informal caregivers. We took the feedback in multiple rounds and incorporated changes along the way. Similar approaches have been conducted successfully by other research teams developing assistive technologies to support self-management of people with dementia [52].

Regarding the digital tablet-based assessment tool, we could incorporate crucial aspects regarding the layout and the handling of the tablet. The difference between our first digital assessment version compared with the final digital assessment version illustrates the importance of involving multiple perspectives, especially of those for whom the assessment instrument or technology is intended. In the literature, there is an increase of studies involving patients and stakeholders in the development or testing of technology [52-54].

The participatory approach allowed us to incorporate aspects into the development of the digital assessment tool that we, as researchers, would have missed otherwise, as can be seen in the first digital tablet-based assessment version in Figure 1 when compared with the final version in Figure 2. The input of informal caregivers, health care professionals, and Alzheimer Society representatives has broadened our perspective and has given us important hints and advice towards aspects that are important for a digital, self-completion unmet needs assessment tool. They identified questions that were unclear, suggested possible answer options, and gave us ideas on how to simplify

the navigation throughout the assessment procedure. Some researchers have used similar procedures for the development of instruments to assess patients' needs [24,47] or to facilitate implementation [19]. A review by Fischer et al [55] reports on the importance of user involvement, particularly the involvement of older adults in technology design. In their review of 40 empirical studies published between 2014 and 2018, these authors examined the consequences of involving older adults in technology design and stated that learning, adjusted design, and improved sense of participation were outcomes frequently stated in studies [55]. Not only did they report that involvement facilitated learning about the needs of older adults but it also led to the generation of new technology ideas such as a companion robot, for instance, which was not in the researchers' mind before involving older adults [55]. Newly gained insights then lead to iterative adjustments of the prototype design [55]. This was also the process that we followed making adjustments to the unmet needs assessment tool. With respect to the quality of technology developed with user involvement, Kopeć and et al [56] noticed that the overall quality of the technology could be improved. They argued that teams who involved users to develop a mobile app were rated higher by an independent jury in a competition than teams who did not. Older adults appreciated the sense of participation and felt that they were being treated as experts on their own lives and equal partners [57]. Researchers reported that involved users enjoyed their experience and some users described their involvement as "happy memories" [58-60].

In contrast to already existing unmet needs assessment instruments, our digital assessment tool identifies a informal caregiver's self-reported unmet needs and ties them directly to interventions most of which are offered within the German health care system. Through the matching process of possible unmet needs with interventions in phase 1 of our assessment development process, we aimed to overcome the pitfalls of identifying problem-centered needs without having support services to meet them at hand. In addition, the digital assessment tool acknowledges the intertwined nature of the dyad, namely the informal caregiver and the person with dementia focusing on not only the health and support needs of the caregiver but also the caregiver's care responsibilities. Although we did not include the CNA-D, we did compare the content and ensured that no important aspect was missed in the assessment procedure.

Limitations

The stakeholders who contributed to the development of the assessment tool were a convenience sample. We have asked informal caregivers, health care professionals and Alzheimer Society representatives who either visited a memory clinic with their significant other, worked at the memory clinic or who were already known to us due to previous collaborations. Furthermore, the sample is limited to people living in a specific region in Northern Germany (Mecklenburg-Western Pomerania). In addition, our level of involvement was limited to stakeholders sharing their feedback on the assessment tool, which is only 1 of the 3 levels of involvement [61]. The tendency to involve stakeholders only in some rather than all levels of involvement was also highlighted by Fischer et al [55] arguing for further research on the actual practices of user involvement. For future projects, a higher level of involvement is sought, especially since current results indicate that an elaborated, theory-based participatory approach including people with dementia, informal caregivers and regional stakeholders might raise the chances of successful implementation in routine care [19].

Conclusion

Our study has shown the feasibility and success of user involvement in the development of an assessment tool for identifying the unmet needs of people with dementia and their informal caregivers. By presenting our approach, we hope to motivate researchers worldwide to expand stakeholder involvement in the early stages of studies. To facilitate this, the questionnaire will be made available to researchers on the homepage of the Network for Translational Dementia Care Research. Further evaluations, such as the acceptance of the tool from the perspective of the study nurses and GPs, are currently being evaluated and will be published separately.

Acknowledgments

The authors would like to acknowledge the input of all stakeholders for their time and perspective. This work was supported by the Federal Joint Committee (G-BA) – Innovation Fund (Gemeinsamer Bundesausschuss/Innovationsausschuss; funding code: 01VSF18030). The ministry had no role in the design of this study and will not have any role during the implementation, analyses, interpretation of the data, or publication of results.

Data Availability

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

Authors' Contributions

Each author has made substantial contributions to this work, has approved the submitted version and has agreed both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Conflicts of Interest

None declared.

References

1. Dementia. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-03-25]
2. Blotenberg I, Hoffmann W, Thyrian JR. Dementia in Germany: epidemiology and prevention potential. Dtsch Arztebl Int 2023 Jul 10;120(27-28):470-476. [doi: 10.3238/arztebl.m2023.0100] [Medline: 37226316]
3. Schulz R, Martire LM. Family caregiving of persons with dementia: prevalence, health effects, and support strategies. Am J Geriatr Psychiatry 2004;12(3):240-249. [doi: 10.1097/00019442-200405000-00002] [Medline: 15126224]
4. Zwingmann I, Hoffmann W, Michalowsky B, et al. Supporting family dementia caregivers: testing the efficacy of dementia care management on multifaceted caregivers' burden. Aging Ment Health 2018 Jul;22(7):889-896. [doi: 10.1080/13607863.2017.1399341] [Medline: 29156941]
5. Michalowsky B, Thyrian JR, Eichler T, et al. Economic analysis of formal care, informal care, and productivity losses in primary care patients who screened positive for dementia in Germany. J Alzheimers Dis 2016;50(1):47-59. [doi: 10.3233/JAD-150600] [Medline: 26639964]
6. Murray J, Schneider J, Banerjee S, Mann A. EUROCARE: A cross-national study of co-resident spouse carers for people with Alzheimer's disease: II—A qualitative analysis of the experience of caregiving. Int J Geriat Psychiatry 1999 Aug;14(8):662-667. [doi: 10.1002/(SICI)1099-1166(199908)14:8<662::AID-GPS993>3.0.CO;2-4]
7. Marin D, Amaya K, Casciano R, et al. Impact of rivastigmine on costs and on time spent in caregiving for families of patients with Alzheimer's disease. Int Psychogeriatr 2003 Dec;15(4):385-398. [doi: 10.1017/s1041610203009633] [Medline: 15000418]
8. Baiyewu O, Smith-Gamble V, Akinbiyi A, et al. Behavioral and caregiver reaction of dementia as measured by the neuropsychiatric inventory in Nigerian community residents. Int Psychogeriatr 2003 Dec;15(4):399-409. [doi: 10.1017/s1041610203009645] [Medline: 15000419]
9. Schneider J, Murray J, Banerjee S, Mann A. EUROCARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: I--Factors associated with carer burden. Int J Geriatr Psychiatry 1999 Aug;14(8):651-661. [doi: 10.1002/(SICI)1099-1166(199908)14:83.0.CO;2-B] [Medline: 10489656]
10. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues Clin Neurosci 2009;11(2):217-228. [doi: 10.31887/DCNS.2009.11.2/hbrodaty] [Medline: 19585957]
11. Berwig M, Heinrich S, Spahlholz J, Hallensleben N, Brähler E, Gertz HJ. Individualized support for informal caregivers of people with dementia – effectiveness of the German adaptation of REACH II. BMC Geriatr 2017 Dec;17(1):286. [doi: 10.1186/s12877-017-0678-y]
12. Williams F, Moghaddam N, Ramsden S, De Boos D. Interventions for reducing levels of burden amongst informal carers of persons with dementia in the community. A systematic review and meta-analysis of randomised controlled trials. Aging Ment Health 2019 Dec 2;23(12):1629-1642. [doi: 10.1080/13607863.2018.1515886]
13. Shata ZN, Amin MR, El-Kady HM, Abu-Nazel MW. Efficacy of A multi-component psychosocial intervention program for caregivers of persons living with neurocognitive disorders, Alexandria, Egypt: A randomized controlled trial. Avicenna J Med 2017;7(2):54-63. [doi: 10.4103/2231-0770.203610] [Medline: 28469987]
14. Lök N, Bademli K. Pilot testing of the "First You Should Get Stronger" program among caregivers of older adults with dementia. Arch Gerontol Geriatr 2017 Jan;68:84-89. [doi: 10.1016/j.archger.2016.09.006]
15. Stirling C, Andrews S, Croft T, Vickers J, Turner P, Robinson A. Measuring dementia carers' unmet need for services - an exploratory mixed method study. BMC Health Serv Res 2010 Dec;10(1):122. [doi: 10.1186/1472-6963-10-122]
16. Bradshaw J, Care M. Taxonomy of social need: problems and progress. In: Medical Care: Essays on Current Research Seventh Series: Nuffield Provincial Hospitals Trust; 1972.
17. Meaney AM, Croke M, Kirby M. Needs assessment in dementia. Int J Geriatr Psychiatry 2005 Apr;20(4):322-329. [doi: 10.1002/gps.1284] [Medline: 15799078]
18. Asadi-Lari M, Tamburini M, Gray D. Patients' needs, satisfaction, and health related quality of life: towards a comprehensive model. Health Qual Life Outcomes 2004 Jun 29;2:32. [doi: 10.1186/1477-7525-2-32] [Medline: 15225377]
19. Seidel K, Quasdorf T, Haberstroh J, Thyrian JR. Adapting a dementia care management intervention for regional implementation: a theory-based participatory barrier analysis. Int J Environ Res Public Health 2022 Apr 30;19(9):5478. [doi: 10.3390/ijerph19095478] [Medline: 35564877]
20. Novais T, Dauphinot V, Krolak-Salmon P, Mouchoux C. How to explore the needs of informal caregivers of individuals with cognitive impairment in Alzheimer's disease or related diseases? A systematic review of quantitative and qualitative studies. BMC Geriatr 2017 Apr 17;17(1):86. [doi: 10.1186/s12877-017-0481-9] [Medline: 28415968]
21. Kipfer S, Pihet S. Reliability, validity and relevance of needs assessment instruments for informal dementia caregivers: a psychometric systematic review. JBI Evid Synth 2020 Apr;18(4):704-742. [doi: 10.11124/JBISRIR-2017-003976] [Medline: 32813339]

22. Bressan V, Visintini C, Palese A. What do family caregivers of people with dementia need? A mixed-method systematic review. Health Soc Care Community 2020 Nov;28(6):1942-1960. [doi: 10.1111/hsc.13048] [Medline: 32542963]
23. Mansfield E, Boyes AW, Bryant J, Sanson-Fisher R. Quantifying the unmet needs of caregivers of people with dementia: a critical review of the quality of measures. Int J Geriatr Psychiatry 2017 Mar;32(3):274-287. [doi: 10.1002/gps.4642] [Medline: 27981694]
24. Wancata J, Krautgartner M, Berner J, et al. The Carers' Needs Assessment for Dementia (CNA-D): development, validity and reliability. Int Psychogeriatr 2005 Sep;17(3):393-406. [doi: 10.1017/s1041610205001699] [Medline: 16252373]
25. Dreier-Wolfgramm A, Michalowsky B, Austrom MG, et al. Dementia care management in primary care: Current collaborative care models and the case for interprofessional education. Z Gerontol Geriatr 2017 May;50(Suppl 2):68-77. [doi: 10.1007/s00391-017-1220-8] [Medline: 28364258]
26. Thyrian JR, Hertel J, Wucherer D, et al. Effectiveness and safety of dementia care management in primary care: a randomized clinical trial. JAMA Psychiatry 2017 Oct 1;74(10):996-1004. [doi: 10.1001/jamapsychiatry.2017.2124] [Medline: 28746708]
27. Zwingmann I, Michalowsky B, Esser A, et al. Identifying unmet needs of family dementia caregivers: results of the baseline assessment of a cluster-randomized controlled intervention trial. J Alzheimers Dis 2019;67(2):527-539. [doi: 10.3233/JAD-180244] [Medline: 30584136]
28. Thyrian JR, Fiß T, Dreier A, et al. Life- and person-centred help in Mecklenburg-Western Pomerania, Germany (DelpHi): study protocol for a randomised controlled trial. Trials 2012 May 10;13(1):56. [doi: 10.1186/1745-6215-13-56] [Medline: 22575023]
29. Dreier A, Thyrian JR, Eichler T, Hoffmann W. Qualifications for nurses for the care of patients with dementia and support to their caregivers: A pilot evaluation of the dementia care management curriculum. Nurse Educ Today 2016 Jan;36:310-317. [doi: 10.1016/j.nedt.2015.07.024] [Medline: 26277428]
30. Thyrian JR, Winter P, Eichler T, et al. Relatives' burden of caring for people screened positive for dementia in primary care : Results of the DelpHi study. Z Gerontol Geriatr 2017 Jan;50(1):4-13. [doi: 10.1007/s00391-016-1119-9] [Medline: 27534949]
31. Eichler T, Thyrian JR, Hertel J, et al. Unmet needs of community-dwelling primary care patients with dementia in Germany: prevalence and correlates. J Alzheimers Dis 2016;51(3):847-855. [doi: 10.3233/JAD-150935] [Medline: 26890767]
32. Köhler L, Meinke-Franze C, Hein J, et al. Does an interdisciplinary network improve dementia care? Results from the IDemUck-study. Curr Alzheimer Res 2014;11(6):538-548. [doi: 10.2174/1567205011666140618100727] [Medline: 24938504]
33. Nikelski A, Keller A, Schumacher-Schönert F, et al. Supporting elderly people with cognitive impairment during and after hospital stays with intersectoral care management: study protocol for a randomized controlled trial. Trials 2019 Aug 30;20(1):543. [doi: 10.1186/s13063-019-3636-5] [Medline: 31470912]
34. Schumacher-Schönert F, Boekholt M, Nikelski A, et al. Versorgungslücken nach dem Krankenhausaufenthalt schließen: Studienergebnisse [intersec-CM] zum Entlass- und Überleitungsmanagement nach § 39 SGB V für Menschen, die im Krankenhaus kognitive demenzielle Beeinträchtigungen zeigen [in German]. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 2024 Apr;185:35-44. [doi: 10.1016/j.zefq.2024.01.001]
35. Gräske J, Meyer S, Schmidt A, et al. Regional dementia care networks in Germany--results from the DemNet-D-Study regarding the quality of life of their users. Pflege 2016;29(2):93-101. [doi: 10.1024/1012-5302/a000477] [Medline: 26974281]
36. Laporte Uribe F, Wolf-Ostermann K, Wübbeler M, Holle B. Care arrangements in dementia care networks: findings from the DemNet-D study baseline and 1-year follow-up. J Aging Health 2018 Jul;30(6):882-903. [doi: 10.1177/0898264317696778] [Medline: 28530156]
37. Zülke A, Luck T, Pabst A, et al. AgeWell.de - study protocol of a pragmatic multi-center cluster-randomized controlled prevention trial against cognitive decline in older primary care patients. BMC Geriatr 2019 Aug 1;19(1):203. [doi: 10.1186/s12877-019-1212-1] [Medline: 31370792]
38. Thyrian JR, Eichler T, Pooch A, et al. Systematic, early identification of dementia and dementia care management are highly appreciated by general physicians in primary care - results within a cluster-randomized-controlled trial (DelpHi). J Multidiscip Healthc 2016;9:183-190. [doi: 10.2147/JMDH.S96055] [Medline: 27143912]
39. Seidl H, Bowles D, Bock JO, et al. FIMA – Fragebogen zur Erhebung von Gesundheitsleistungen im Alter: entwicklung und pilotstudie [Article in German]. Gesundheitswesen 2015;77(1):46-52 [FREE Full text] [doi: 10.1055/s-0034-1372618]
40. Higginson IJ, Gao W, Jackson D, Murray J, Harding R. Short-form Zarit Caregiver Burden Interviews were valid in advanced conditions. J Clin Epidemiol 2010 May;63(5):535-542. [doi: 10.1016/j.jclinepi.2009.06.014] [Medline: 19836205]
41. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503]
42. Zank S, Schacke C, Leipold B. Berliner Inventar zur Angehörigenbelastung - Demenz (BIZA-D). Z Klin Psychol Psychother 2006 Oct;35(4):296-305. [doi: 10.1026/1616-3443.35.4.296]
43. Reynolds T, Thornicroft G, Abas M, et al. Camberwell Assessment of Need for the Elderly (CANE). development, validity and reliability. Br J Psychiatry 2000 May;176(5):444-452. [doi: 10.1192/bjp.176.5.444] [Medline: 10912220]

44. Oliveira DC, Vass C, Aubeeluck A. The development and validation of the Dementia Quality of Life Scale for Older Family Carers (DQoL-OC). Aging Ment Health 2018 May;22(5):709-716. [doi: 10.1080/13607863.2017.1293004] [Medline: 28282723]
45. Monahan PO, Alder CA, Khan BA, Stump T, Boustani MA. The Healthy Aging Brain Care (HABC) Monitor: validation of the Patient Self-Report Version of the clinical tool designed to measure and monitor cognitive, functional, and psychological health. Clin Interv Aging 2014;9:2123-2132. [doi: 10.2147/CIA.S64140] [Medline: 25584024]
46. Stein J, Pabst A, Weyerer S, et al. The assessment of met and unmet care needs in the oldest old with and without depression using the Camberwell Assessment of Need for the Elderly (CANE): Results of the AgeMooDe study. J Affect Disord 2016 Mar 15;193:309-317. [doi: 10.1016/j.jad.2015.12.044] [Medline: 26774519]
47. King RB, Hartke RJ, Lee J, Raad J. The stroke caregiver unmet resource needs scale: development and psychometric testing. J Neurosci Nurs 2013 Dec;45(6):320-328. [doi: 10.1097/JNN.0b013e3182a3ce40] [Medline: 24217142]
48. EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
49. Gray J, Kim J, Ciesla JR, Yao P. Rasch analysis of the Lubben Social Network Scale-6 (LSNS-6). J Appl Gerontol 2016 May;35(5):508-528. [doi: 10.1177/0733464814560468] [Medline: 25428591]
50. Hébert R, Bravo G, Préville M. Reliability, validity and reference values of the Zarit Burden Interview for assessing informal caregivers of community-dwelling older persons with dementia. Can J Aging 2000;19(4):494-507. [doi: 10.1017/S0714980800012484]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
52. Øksnebjerg L, Woods B, Ruth K, et al. A tablet app supporting self-management for people with dementia: explorative study of adoption and use patterns. JMIR Mhealth Uhealth 2020 Jan 17;8(1):e14694. [doi: 10.2196/14694] [Medline: 31951217]
53. Hettinga M, De Boer J, Goldberg E, Moelaert F. Navigation for people with mild dementia. Stud Health Technol Inform 2009;150:428-432. [doi: 10.3233/978-1-60750-044-5-428] [Medline: 19745347]
54. van de Vijver S, Hummel D, van Dijk AH, et al. Evaluation of a digital self-management platform for patients with chronic illness in primary care: qualitative study of stakeholders' perspectives. JMIR Form Res 2022 Aug 3;6(8):e38424. [doi: 10.2196/38424] [Medline: 35921145]
55. Fischer B, Peine A, Östlund B. The importance of user involvement: a systematic review of involving older users in technology design. Gerontologist 2020 Sep 15;60(7):e513-e523. [doi: 10.1093/geront/gnz163] [Medline: 31773145]
56. Kopeć W, Balcerzak B, Nielek R, Kowalik G, Wierzbicki A, Casati F. Older adults and hackathons: a qualitative study. Empir Software Eng 2018 Aug;23(4):1895-1930. [doi: 10.1007/s10664-017-9565-6]
57. Maaß S, Buchmüller S. The crucial role of cultural probes in participatory design for and with older adults. i-com 2018 Aug 28;17(2):119-135. [doi: 10.1515/icom-2018-0015]
58. Eftring H, Frennert S. Designing a social and assistive robot for seniors. Z Gerontol Geriatr 2016 Jun;49(4):274-281. [doi: 10.1007/s00391-016-1064-7] [Medline: 27220732]
59. Lee HR, Šabanović S, Chang WL, et al. Steps toward participatory design of social robots: mutual learning with older adults with depression. Presented at: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction; Mar 6-9, 2017; Vienna Austria p. 244-253. [doi: 10.1145/2909824.3020237]
60. Alaoui M, Ting KLH, Lewkowicz M. The urge for empirically-informed design of social-oriented AAL applications – the example of 2 AAL projects. Procedia Comput Sci 2014;27:379-388. [doi: 10.1016/j.procs.2014.02.042]
61. Arnstein SR. A ladder of citizen participation. J Am Inst Plann 1969 Jul;35(4):216-224. [doi: 10.1080/01944366908977225]

Abbreviations

CANE: Camberwell Assessment of Need for the Elderly
CNA-D: Carer's Needs Assessment for Dementia
DelpHi-MV: Dementia: life- and person-centered help in Mecklenburg-Western Pomerania
G-BA: Gemeinsamer Bundesausschuss/Innovationsausschuss
GAIN: Gesund Angehörige Pflegen
GP: general practitioner
HCP: health care professional
JPND: EU Joint Programme in Neurodegenerative Research
LSNS: Lubben Social Network Scale
VAS: visual analog scale
WHO: World Health Organization
ZBI: Zarit Burden Interview

Edited by R Bjarnadottir; submitted 26.04.24; peer-reviewed by C Peterson, M Patrascu; revised version received 02.12.24; accepted 23.12.24; published 07.04.25.

Please cite as:
Biernetzky OA, Thyrian JR, Boekholt M, Berndt M, Hoffmann W, Teipel SJ, Kilimann I
Identifying Unmet Needs of Informal Dementia Caregivers in Clinical Practice: User-Centered Development of a Digital Assessment Tool
JMIR Aging 2025;8:e59942
URL: https://aging.jmir.org/2025/1/e59942
doi:10.2196/59942

Uncovering Specific Navigation Patterns by Assessing User Engagement of People With Dementia and Family Caregivers With an Advance Care Planning Website: Quantitative Analysis of Web Log Data

Charlèss Dupont[1,2], PhD; Tinne Smets[1,2], PhD; Courtney Potts[3], MSc; Fanny Monnet[1,2], PhD; Lara Pivodic[1,2], PhD; Aline De Vleminck[1,2], PhD; Chantal Van Audenhove[4], PhD; Maurice Mulvenna[5], PhD; Lieve Van den Block[1,2], PhD

[1]End-of-life Care Research Group, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels, Belgium

[2]Department of Family Medicine and Chronic Care, Vrije Universiteit Brussel (VUB), Brussels, Belgium

[3]School of Psychology, Ulster University, Coleraine, United Kingdom

[4]LUCAS Center for Care Research and Consultancy, KU Leuven, Leuven, Belgium

[5]School of Computing, Ulster University, Belfast, United Kingdom

Corresponding Author:
Charlèss Dupont, PhD
End-of-life Care Research Group, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, Brussels, Belgium

Abstract

Background: Web-based tools have gained popularity to inform and empower individuals in advance care planning. We have developed an interactive website tailored to the unique needs of people with dementia and their families to support advance care planning. This website aims to break away from the rigid pathways shown in other tools that support advance care planning, in which advance care planning is shown as a linear process from information to reflection, communication, and documentation.

Objective: This study aimed to assess the website's usage by people with dementia and their family caregivers, identify distinct user engagement patterns, and visualize how users navigated the website.

Methods: We analyzed the website's log data obtained from an 8-week evaluation study of the site. Interactions with the website were collected in log data files and included visited web pages or clicked-on hyperlinks. Distinct user engagement patterns were identified using K-means clustering process mining, a technique that extracts insights from log data to model and visualize workflows, was applied to visualize user pathways through the website.

Results: A total of 52 participants, 21 individuals with dementia and their family caregivers as dyads and 10 family caregivers were included in the study. Throughout the 8-week study, users spent an average of 35.3 (SD 82.9) minutes over 5.5 (SD 3.4) unique days on the website. Family caregivers mostly used the website (alone or with a person with dementia) throughout the 8-week study. Only 3 people with dementia used it on their own. In total, 3 distinct engagement patterns emerged: low, moderate, and high. Low-engagement participants spent less time on the website during the 8 weeks, following a linear path from information to communication to documentation. Moderate- and high-engagement users showed more dynamic patterns, frequently navigating between information pages and communication tools to facilitate exploration of aspects related to advance care planning.

Conclusions: The diverse engagement patterns underscore the need for personalized support in advance care planning and challenge the conventional linear advance care planning representations found in other web-based tools.

(JMIR Aging 2025;8:e60652) doi:10.2196/60652

KEYWORDS

dementia; advance care planning; user engagement; web-based tool; care; website; caregiver; communication; tool; online

Introduction

Advance care planning is a communication process between patients, families, and health care professionals to "define goals and preferences for future care and treatment" [1]. This process holds particular relevance in dementia [2]. The cognitive decline in dementia highlights the need for early initiation of advance care planning, enabling people with dementia to reflect on and express their preferences for future care and treatments [3-5]. While existing definitions of advance care planning often focus on medical care decisions [6], people with dementia and their families have emphasized that it should include exploring what matters now and in the future, including nonmedical aspects of care [7-10]. Furthermore, people with dementia and their

families have expressed a need to discuss future care together [7]. This aligns with the recently introduced public health to advance care planning, emphasizing a shift toward a social focus on "what matters most to people" rather than the current emphasis on end-of-life decision making and underlining the need to support conversations in the family context [11].

To support advance care planning within the family context, interactive web-based tools like websites or apps have been promoted [12]; however, despite their proven benefits in other populations, there is a noticeable absence of tools tailored to the specific needs of people with dementia and their families [13-15]. To address this gap, we developed a website for and with people with dementia and their families [16-19]. This website deviates from the structured linear pathways found in other tools to support advance care planning [15]. Such tools adhere to a stepwise procedure, often commencing with information provision, prompting reflection, followed by communication, and concluding with documentation, commonly in the form of advance directives. For example, in some cases, users must complete or make an explicit effort to skip a step before being able to proceed to the next step, eg, users cannot access certain features such as communication tools or documentation forms without first completing previous tasks, for example, users have to follow a linear pathway when navigating the website [15,20]. In contrast, our website is designed to offer flexibility, allowing users to engage with the content in any way they choose. There are no predefined steps or sequences to follow, and users can access all features, such as communication tools and documentation forms, without needing to complete any previous tasks [16].

Following the development of the website to support advance care planning, we performed an 8-week evaluation study involving people with mild to moderate dementia and their family caregivers [17,18]. The primary focus of this evaluation was to assess the usability, acceptability, feasibility, experiences, and outcomes of using the website [17,18]. Participants in the evaluation study found that the website supports advance care planning [17,18]. After 8 weeks, participants exhibited improved advance care planning knowledge, self-efficacy, and skills [17,18].

Beyond this evaluation, gaining insights into the website's user engagement and usage patterns is crucial to enhance our understanding of how people with dementia and their families use a website to support advance care planning in the family context, how they engage with it and how single users and dyadic users differ (ie, people with dementia together with their family caregiver). This knowledge will inform enhancements to the website's design and functionality by highlighting user needs in digital health interventions, ultimately improving support for families navigating advance care planning. Therefore, this study aims to: explore how people with dementia and their family caregivers used the advance care planning support website during an 8-week evaluation study; explore whether and which user behavior clusters can be typified based on the engagement of people with dementia and their family caregivers with the website; and explore and visualize user pathways of the identified user behavior clusters, that is, how the different user clusters of people with dementia and their family caregivers navigate through the website.

Methods

Study design

This study quantitatively analyzed web log data from a convergent parallel mixed methods evaluation study of a website designed to support people with dementia and their family caregivers in advance care planning. Web log data was collected during an 8-week evaluation study to capture participants' interactions with the website, for example, time spent on the website and what pages were visited. This log data was used to explore the website's usage by people with dementia and their family caregivers, which user behavior clusters can be typified and to visualize user pathways. The protocol, including the objective to explore usage patterns of the website, as well as the main results of the evaluation study, have been published elsewhere [17,18]

The Development of the Advance Care Planning Support Website

The development of the website followed a user-centred, iterative design process, ensuring alignment with the needs of people with dementia and their families. The website was developed to provide information and support for people with dementia and their family caregivers in advance care planning [16]. Furthermore, the website was designed with flexible navigation options, allowing users to explore different sections at their own pace rather than following a linear, step-by-step approach. This was particularly important given the varying levels of readiness among users to engage with sensitive topics such as future care preferences [8,16].

The website includes advance care planning information, information about legal frameworks, communication tips, and documentation sections. Furthermore, it provides accessibility features, such as text-to-speech and text enlargement and 2 interactive communication tools, an "Interactive Card Tool" based on the recently developed paper-based version of the Levenswensen (Life Wishes) cards and a fill-in tool "Thinking Now About Later," that guides users through a reflective process to help users think and talk about and write down their preferences for the present and future (Multimedia Appendices 1 and 2). The results of the evaluation study and experiences with the interactive communication tools, have been published elsewhere [16,18,21].

Participants, Recruitment and Setting

People with mild to moderate dementia, including both early and late onset, along with their family caregivers as dyads or the family caregiver alone, were recruited [17]. To be eligible, participants needed to meet the criteria shown in Textbox 1 [17]
.

Textbox 1. Inclusion criteria for the study.

Both people with dementia and their family caregivers:

- Express interest in and willingness to test the website to support advance care planning.
- Provide consent for study participation.
- Proficient in speaking and understanding Dutch.
- Own a device (laptop, tablet, mobile phone, etc).

People with dementia:

- Diagnosed with young- or late-onset dementia.

Family caregivers:

- Actively involved in the care (physical, emotional, social, etc) of the person with dementia.

For participants recruited as dyads, at least one had to be capable of navigating the website. For instance, the person with dementia and the family caregiver could not both have visual impairments or other disabilities hindering interaction with the advance care planning website. Participants were recruited through dementia care organizations and memory clinic neurologists. Eligible participants with a confirmed dementia diagnosis were identified by health care professionals. People with dementia who were interested in the study contacted the researcher, received study information, and, if still interested after receiving the study information, were offered an appointment to perform the eligibility screening and provide informed consent. The recruitment process is described in more detail in the study protocol that is published elsewhere [17]. The study was conducted in Flanders, the Dutch-speaking region of Belgium.

Data Collection Procedures

Between October 2022 and May 2023, 52 people participated in the evaluation study: 21 dyads, that is, 21 people with dementia and 21 family caregivers, and 10 family caregivers alone. Information on sociodemographic data, including age, gender, computer literacy, type of diagnosis and date, was gathered through a survey assessed at the start of the 8 weeks [17,18]. Participants were granted access to the website after providing informed consent and completing baseline data collection. They were informed that they could use the website freely over an 8-week period, choosing how and when to engage with it according to their preferences [17].

To capture the usage by the participants during the 8-week study, there was a continuous collection of log data. Usage refers to how people with dementia and family caregivers engage with the advance care planning website. The log data captured what pages were visited, time spent on each page, interaction with content or functions, frequency of visits to each page, and the search queries used. Upon accessing the website, participants were prompted to identify themselves through a pop-up question, requiring them to fill in their study ID and specify whether they were a person with dementia, a family caregiver, or engaging together as a dyad. A detailed study protocol of the evaluation study is published elsewhere [8].

Data Analysis

Overview

Log data of the advance care planning website were analyzed using the programming language R (version 4.2.3, R Studio). The log data was saved in 3 files, application and access logs and a file to save the interactions with the interactive communication tools. First, the 3 log datasets were cleaned by eliminating irrelevant data, such as admin and php requests. The application logs were further filtered to include only valid study IDs used by the participants. Study IDs were then added to the access logs based on cross matching with IP addresses. The access and application logs were combined, and additional information from the third dataset regarding the completion of interactive communication tools was incorporated by matching IP addresses to study IDs.

Assessing the Extent to Which the Website Was Used

First, the interactions with the website were summarized by time spent on the website, pages visited, and who visited. This analysis focused on usage patterns, which provided an initial understanding of the variability in engagement.

Identifying Distinct Behavioral Clusters

After, the data underwent K-means clustering. Clustering the data aimed to identify behavioral or interaction patterns that typify user engagement. The 6 features derived from the summary data were used for clustering: total interactions, unique days, duration of use, total clicks on communication, information, and documentation pages. The R package “caret” was used to normalize the data using min-max normalization, converting variables to a range between 0 and 1 [22]. Usage frequency emerged as the key differentiating factor among participants. The “NbClust” package was then used to compare summary statistics and determine the optimal number of clusters, using 30 indices to determine the optimal number of clusters (k) between 2 and 10 based on these metrics [23]. The “NbClust” package indicated that the data could be best classified into 3 clusters, which were named after usage frequency of participants, for example , low, moderate, and high engagement levels. Subsequently, the k-means algorithm was applied to the data for k clusters. Principal component analysis (PCA) was used for data visualization to reduce the multiple features to 2 dimensions. Finally, Kruskal-Wallis rank sum tests

and χ^2 tests were conducted to evaluate the significant differences between features for each cluster, providing further insights into the characteristics of the identified user clusters.

Identification of User Pathways

To visualize how users navigated through the website, we applied process mining techniques using the R package BupaR. This allowed us to map and analyze the various paths users followed across the site's different sections [24]. Log data were filtered to remove redundant information (taken out: change font size, contrast, privacy policy, read speaker, print, and other). First, a process matrix, which is a 2-dimensional matrix showing the flows between activities, was generated to visualize the entire log dataset [25]. After, individual process maps were produced per study ID to show the path taken by participants. Finally, the individual paths were compared with the IDs per cluster to find similarities and differences between the participants' paths in each cluster.

Ethical Considerations

The research protocol was submitted to the Medical Ethics Committee of Brussels university hospital (UZ Brussel) and received ethical approval (BUN 1432022000179). All participants provided informed consent, and the study was conducted in accordance with the relevant ethical guidelines and regulations. To ensure privacy and confidentiality, all participant data were de-identified prior to analysis. Protective measures, such as secure data storage and restricted access to personal information, were implemented to safeguard participant privacy throughout the study.

Results

Overview

In total, 52 participants with 31 study IDs were included (21 dyads and 10 family caregivers), consisting of 21 people with dementia and 31 family caregivers. Ten family caregivers participated alone, either because the person with dementia was unable to provide informed consent (6/10) or chose not to participate (4/10). The average age of people with dementia was 62.8 (SD 10.4) years, with 42.8% (9/21) being female. Among family caregivers, 68% (21/31) were female, with an average age of 62.1 (SD 10.9) years. Family caregivers reported an average computer literacy score of 7.5 (on a scale from 1 to 10, with higher scores indicating greater self-reported computer literacy), while people with dementia had an average score of 4 (Table 1).

Table . Description of the study population.

Variables		Persons with dementia (n=21)	Family caregivers (n=31)
Sex (female), n		9	21
Age (years), mean (SD)		62.1 (10.9)	62.8 (10.4)
Relationship, n			
	Partners[a]	18	25
	Parent (in law)-child	3	6
Profession, n			
	Employed	1	16
	Retired	20[b]	15
Computer literacy[c], mean (SD)		62.1 (10.9)	62.8 (10.4)
Dementia diagnosis[d], n			
	Alzheimer disease	15	20
	Vascular dementia	1	3
	Frontotemporal dementia	3	3
	Lewy body dementia	1	1
	Parkinson dementia	0	1
	I do not know	1	3

[a]Married, living together or in a romantic relationship.

[b]Six of the retired persons with dementia were forced to take early retirement because of their diagnoses.

[c]Self-evaluation on a scale from 1 (no computer skills) to 10 (excellent computer skills).

[d]Dates of diagnoses: January 2013 to December 2022.

Website Usage

Each study ID (n=31) was logged at least once in the log data, indicating at least 1 visit by 1 or both dyad members and by each family caregiver who participated alone. In 10 study IDs, both the family caregiver and dyad used the website. In 15 instances, only the caregiver accessed it and in 3 cases, solely the dyad engaged with the website. In addition, 3 occurrences

involved mixed usage, with indications of the family caregiver, dyad, or person with dementia using the website. The total number of unique interactions by all users (n=31) with the website was 1799 (ie, the total interaction data points), encompassing information searches, clicks, and use of interactive elements. On average, users had 58 (SD 57) interactions over the 8 weeks. The total duration spent on the website during the 8 weeks was 35.3 (SD 82.9) minutes, and, on average, people used the website on 5.5 (SD 3.4) unique days of the 8-week study period.

User interactions revealed that family caregivers had the highest overall number of interactions (757), followed by dyads (235) and people with dementia alone (103). We faced difficulty attributing the information to specific user types for the other 701 interactions documented in the access log file. These unidentified usages per user type arose from difficulties linking interactions with corresponding user types when users left their browsers open for extended periods. The webpage "Advance Care Planning: What Is It?" was visited most (n=304). Followed by the glossary section (209), "Advance Care Planning: Thinking and Talking About Later" (277), "Advance Care Planning: Writing It Down for Later" (259) and the frequently asked questions section (122). The 2 interactive communication tools were used 136 (Thinking Now About Later) and 91 (Levenswensen cards) times.

Identifying Typical User Behaviors by Clustering

To determine the optimal number of clusters, we applied NbClust with 30 indices, revealing that 8 favored 2 clusters and 8 indicated 3 clusters. Following the majority rule, we opted for 3 clusters, each assigned to study IDs reflecting diverse user engagement patterns. Table 2 summarizes each cluster's characteristics. To showcase the most-visited web pages by the clusters, we categorized all pages into 10 categories (Figure 1; Multimedia Appendix 1 visually shows the web pages corresponding to the categories).

Table . Overview of the characteristics of use of the website of each cluster.

Characteristics	Moderate engagement level	Low engagement level	High engagement level	*P* value
Size, study ID/n[a] (%)	15/21 (45)	5/10 (21)	11/16 (34)	—[b]
Total interactions[c], mean (SD)	50 (13)	21 (9)	86 (15)	<0.001[d]
Unique days, mean (SD)	4.4 (1.5)	1.6 (0.5)	8.7 (3.3)	0.004[d]
Range of use (days), mean (SD)[e]	56.7 (19.3)	9.8 (19.2)	65.6 (16.9)	<0.001[d]
Communication pages visited, mean (SD)	17 (7)	5 (3)	28 (10)	<0.001[d]
Documentation pages visited, mean (SD)	7 (4)	6 (4)	18 (4)	<0.001[d]
Information pages visited, mean (SD)	22 (10)	7 (5)	34 (13)	<0.001[d]
Total number of interactions per user type, n				<0.001[g]
Caregiver	420	27	310	
Person with dementia	0	0	103	
Dyad	124	33	88	
Family caregiver and person with dementia[f]	204	54	452	

[a]n is the number of users (N=46 because 5 people with dementia did not use the website or participate anymore).

[b]Not applicable.

[c]Interaction is any movement of the user on the website. This can be clicking on a hyperlink, watching a video, opening a page, or printing the web page.

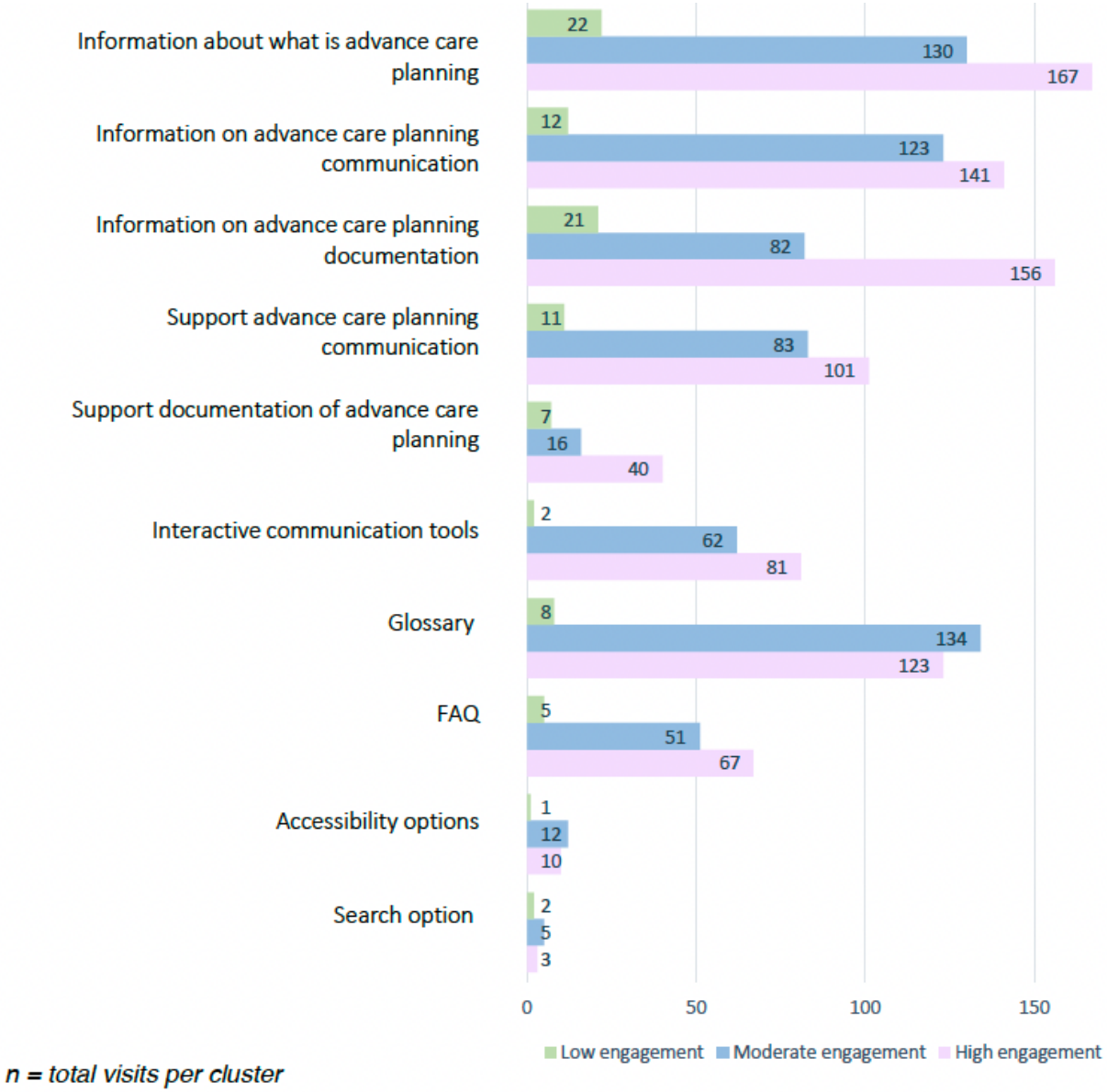
[d]χ^2 test.

[e]From the first day of use until the last day. For example, cluster moderate engagement started using the website, and, on average, the last usage was 9.8 days later.

[f]When the IP address could not be matched with the date or time of the filled-in study ID.

[g]Kruskal-Wallis rank sum test.

Figure 1. Overview of the most-visited web pages by the 3 clusters. FAQ: frequently asked questions.



Cluster low engagement (5 participants) exhibited the fewest interactions (mean of 21), shorter duration (mean of 9.8 days), and accessed fewer pages. Cluster moderate engagement (15 participants) had a mean of 50 interactions, visited for 4.4 unique days, with a use duration of approximately 56.7 days (Table 2). Cluster high engagement (11 participants) demonstrated the highest engagement, with the highest total interactions (mean of 86), longer duration (mean of 65.6 days), accessing more diverse pages and including the 3 people with dementia who used the website alone. Statistically significant *P* values underscore distinct engagement patterns among the clusters (Table 2).

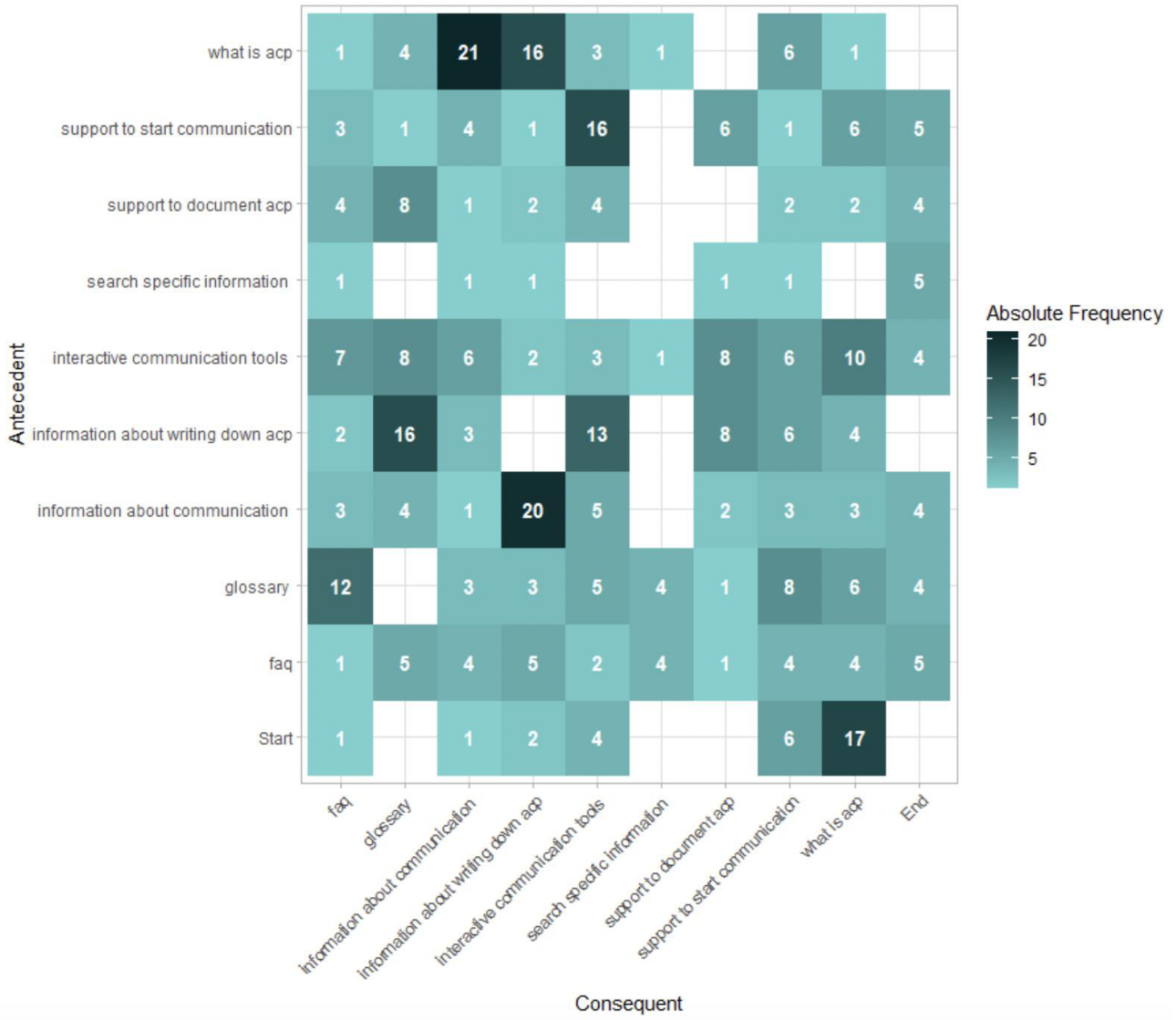
Participants across all 3 clusters predominantly visited the "what is advance care planning" web page the most. Low-engagement participants mostly focused on the information pages, especially the documentation information and showed limited interest in interactive tools. Moderate engagement participants visited mostly the advance care planning information pages, frequently using the glossary and accessibility features like contrast and text-to-speech, distinguishing their usage from other clusters. In the high-engagement cluster, participants explored information, communication and documentation pages, actively using interactive communication tools.

Identification of User Pathways of the Identified User Behavior

To identify the user pathways, we first look at the overall pathways of users using a process matrix (Figure 2). The process matrix is a two-dimensional representation that illustrates the flow between the web pages that users have visited. The matrix is organized with antecedent events followed by the consequent events. The primary pathway, observed in 21 instances, initiates with a visit to the "What is advance care planning" page, followed by navigating to a subsequent page providing "information about communication." Significantly, users frequently started their pathway on informative pages, such as "what is advance care planning," "information on communication," or "Information on documentation." Subsequently, they progress to explore additional information pages. Noteworthy is the observation that users typically visit

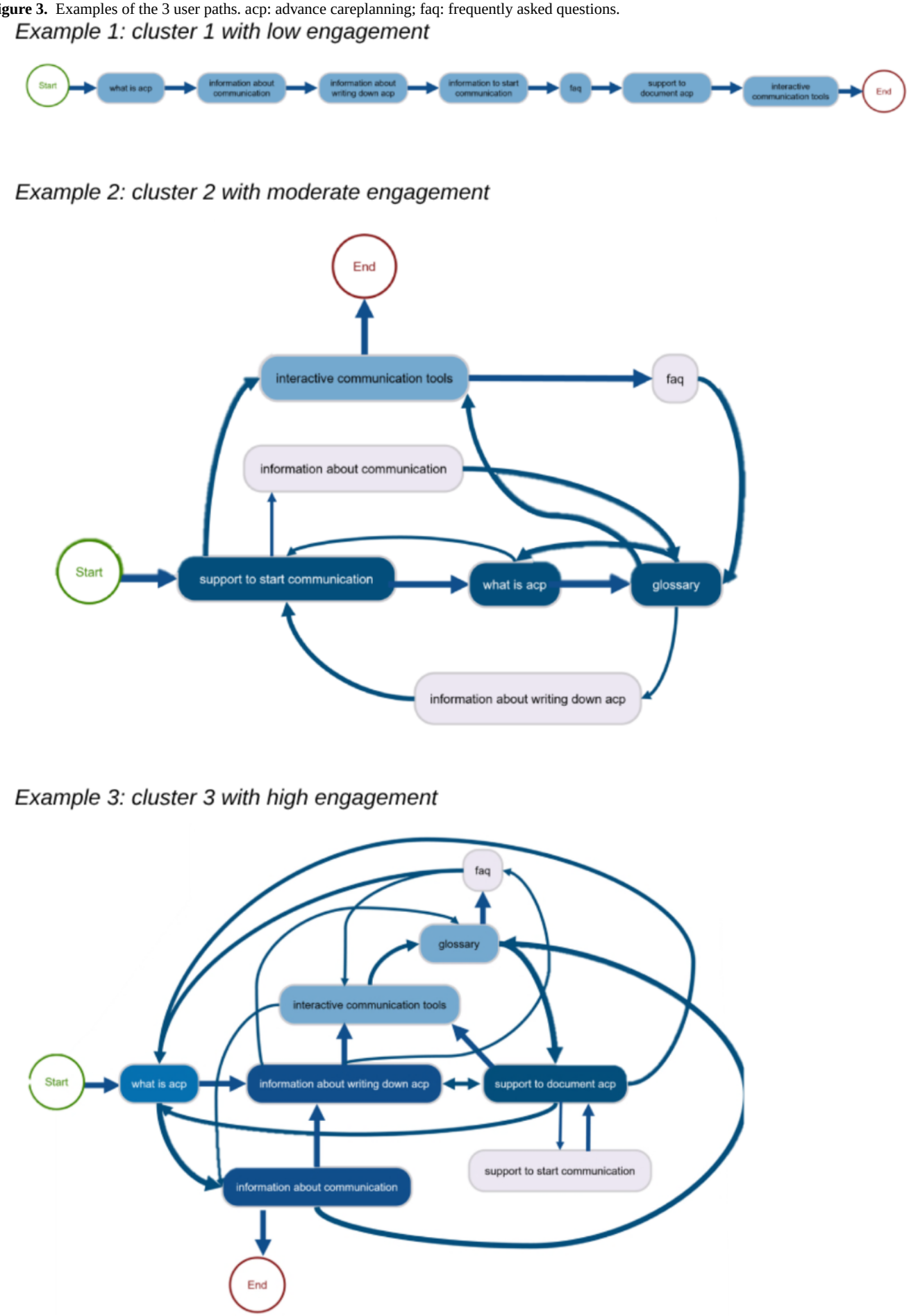
pages providing information about communication or documentation before engaging with the interactive communication tools.

Figure 2. Identification of overall user pathways. acp: advance care planning; FAQ: frequently asked questions.



Then, the user pathways of the identified user behavior clusters were identified (Figure 3). Low-engagement participants (Figure 3, example 1) displayed a linear and direct browsing style, rarely revisiting previous pages during navigation. In contrast, the participants with moderate and high engagement (Figure 3, examples 2 and 3) explored the website by visiting pages sequentially, occasionally revisiting previously viewed pages. High engagement extensively explored various pages, moving between information and guidance pages. Participants with moderate engagement involved frequent transitions between pages but less often than those with high engagement.

Figure 3. Examples of the 3 user paths. acp: advance careplanning; faq: frequently asked questions.



Discussion

Principal Findings

The analyses of the log data of an 8-week evaluation study of a website to support people with dementia and their families in advance care planning showed that family caregivers used the website most often, either alone or with a person with dementia, that is, dyad. Three distinct engagement patterns emerged in this study: low, moderate, and high. Low-engagement users tended to follow a more linear path on the website, while moderate- and high-engagement users displayed a more dynamic engagement, exploring the website in diverse ways.

Flexible user navigation patterns were evident in our study, challenging the conventional linear advance care planning representations found in other web-based tools in which users go through a stepwise process, typically starting with information provision, prompting reflection, moving on to communication, and concluding with documentation in the form of advance directives [15]. This rigid structure, embedded in the design of other ACP websites, may restrict easy access to subsequent features or content until earlier steps are completed. While this structured pathway might work for some users, it assumes that everyone is ready to follow the same route through the ACP process, which may not always be the case [7,8,26]. In our study, low-engagement users tended to follow a more linear pathway. However, those with moderate or high engagement demonstrated more dynamic and varied usage patterns. Some users initiated their engagement by seeking information, while others prioritized documentation or directly accessed communication tools. This flexibility of navigation aligns with feedback from family caregivers during the website's development process, which emphasized the importance of allowing users to navigate at their own pace, rather than following a rigid, linear path [8]. This flexibility allowed for personalized engagement, as different users may be at different stages of readiness to engage with sensitive topics, such as future care planning. Furthermore, this approach aligns with broader technological research [7,10,26], which shows that family caregivers prefer flexible navigation in internet-based tools [27,28]. These findings suggest a need for a more flexible approach, indicating that users should have the freedom to navigate tools that align best with their needs, which is also supported by other research [5,26,29]. This also aligns with Belgian clinical guidelines [30], explicitly mentioning that health care professionals should tailor advance care planning in dementia, including style and content, to the "person's level and rhythm."

Despite participants' expressed interest in advance care planning in the evaluation study, people with dementia rarely engaged with the website on their own. Family caregivers and the person with dementia did engage together, emphasizing the family's importance. This finding is not necessarily surprising as much literature points at the importance of family in a dementia trajectory. A recently published consensus definition on advance care planning in dementia also highlighted family as highly important and specific in this population [6,31]. Regarding the use or uptake of websites among people with dementia, involving family caregivers can play a facilitating role; however, it is essential to acknowledge that not all people with dementia have family caregivers or families directly engaged in their care [32]. Furthermore, overreliance on family caregivers may unintentionally hinder independent usage, potentially undermining autonomy. Therefore, achieving an inclusive environment necessitates balancing involving family caregivers and promoting self-usage.

Strengths and Limitations of the Study

This study has several strengths that contribute to the robustness of our study's findings. Using log data, this study is the first to examine user engagement of people with dementia and their families with a website to support advance care planning. It offers a comprehensive understanding of their specific usage patterns. The elimination of recall bias is another key strength, as log data provided an accurate account of how users engaged with the website. The study also has limitations with regard to the data used. We encountered difficulty identifying the specific type of user for all log data because not all access logs could be matched with the date or time of the filled-in type of user in the application log. In addition, due to a 1-month retention period for log data, a small portion of data (7 days with interactions from 2 users, as indicated in Google Analytics) was lost as it was not downloaded before deletion from the server. Another limitation is the relatively small dataset for our log data analyses. Specifically, the number of participants with dementia who accessed the website independently was small what limits our ability to make comparisons in usage patterns between people with dementia, family caregivers, and dyads. Further research with a larger sample of people with dementia, including both independent and caregiver-supported users, is needed to enhance the robustness of the findings and provide stronger insights into their engagement patterns. Finally, a limitation arises from self-selection bias in user type; for example, users might identify as "family caregivers" while engaging together, and vice versa, introducing variability that could impact the accuracy of findings.

Conclusions

This study offers insights in how people with dementia and their family caregivers use a website designed to support them in advance care planning. The findings show that family caregivers are the website's primary users, often engaging with the website alone or together with the person with dementia. Three distinct user engagement patterns, low, moderate, and high engagement, were identified, with more dynamic navigation observed among high-engagement users, particularly those using interactive communication tools. These findings underscore the need for flexible user pathways in online advance care planning tools, allowing personalized navigation. Future research with a larger and more diverse sample of persons with dementia is necessary to confirm these findings and allow for more detailed comparisons across different user groups.

Acknowledgments

We would like to thank all participants for their time and effort. We thank the neurologists working in memory clinics, the Expertise Centrum Dementia Flanders, Alzheimer Liga Flanders, and Het Ventiel for their efforts in recruiting participants. This study is part of the project "CAPACITY: Flanders Project to Develop Capacity in Palliative Care Across Society," a collaboration between the Vrije Universiteit Brussel, Ghent University, and the Catholic University Leuven, Belgium. This study is supported by a grant from the Research Foundation – Flanders (file number S002219N). The research presented in this paper was carried out as part of the DISTINCT project, a Marie Curie Innovative Training Network (ITN), H2020-MSCA-ITN-2018, under grant agreement number 813196.

Data Availability

The website assessed in this study is available in Dutch [19].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the website.
[DOCX File, 4672 KB - aging_v8i1e60652_app1.docx]

Multimedia Appendix 2
The interactive communication tools.
[DOCX File, 1664 KB - aging_v8i1e60652_app2.docx]

References

1. Rietjens JAC, Sudore RL, Connolly M, et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. Lancet Oncol 2017 Sep;18(9):e543-e551. [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
2. Van den Block L. Advancing research on advance care planning in dementia. Palliat Med 2019 Mar;33(3):259-261. [doi: 10.1177/0269216319826411] [Medline: 30758274]
3. Bosisio F, Jox RJ, Jones L, Rubli Truchard E. Planning ahead with dementia: what role can advance care planning play? A review on opportunities and challenges. Swiss Med Wkly 2018 Dec 17;148:w14706. [doi: 10.4414/smw.2018.14706] [Medline: 30594990]
4. van der Steen JT, Radbruch L, Hertogh CMPM, et al. White paper defining optimal palliative care in older people with dementia: a Delphi study and recommendations from the European Association for Palliative Care. Palliat Med 2014 Mar;28(3):197-209. [doi: 10.1177/0269216313493685] [Medline: 23828874]
5. Wendrich-van Dael A, Bunn F, Lynch J, Pivodic L, Van den Block L, Goodman C. Advance care planning for people living with dementia: an umbrella review of effectiveness and experiences. Int J Nurs Stud 2020 Jul;107:103576. [doi: 10.1016/j.ijnurstu.2020.103576] [Medline: 32380259]
6. van der Steen JT, Nakanishi M, Van den Block L, et al. Consensus definition of advance care planning in dementia: a 33-country Delphi study. Alzheimers Dement 2024 Feb;20(2):1309-1320. [doi: 10.1002/alz.13526] [Medline: 37985444]
7. Dickinson C, Bamford C, Exley C, Emmett C, Hughes J, Robinson L. Planning for tomorrow whilst living for today: the views of people with dementia and their families on advance care planning. Int Psychogeriatr 2013 Dec;25(12):2011-2021. [doi: 10.1017/S1041610213001531] [Medline: 24053783]
8. Dupont C, Smets T, Monnet F, et al. Defining the content of a website on advance care planning in dementia: a focus group study with family and health professionals. BMC Med Inform Decis Mak 2023 Nov 8;23(1):254. [doi: 10.1186/s12911-023-02359-1] [Medline: 37940991]
9. Monnet F, Pivodic L, Dupont C, Diaz A, Gove D, Van den Block L. Defining advance care planning from the perspective of people with dementia: focus groups with the European Working Group of People with dementia. Alzheimers Dement 2023 Jun;19(S5). [doi: 10.1002/alz.063913]
10. Van Rickstal R, De Vleminck A, Aldridge MD, et al. Limited engagement in, yet clear preferences for advance care planning in young-onset dementia: an exploratory interview-study with family caregivers. Palliat Med 2019 Oct;33(9):1166-1175. [doi: 10.1177/0269216319864777] [Medline: 31342861]
11. Abel J, Kellehear A, Millington Sanders C, Taubert M, Kingston H. Advance care planning re-imagined: a needed shift for COVID times and beyond. Palliat Care Soc Pract 2020;14:2632352420934491. [doi: 10.1177/2632352420934491] [Medline: 32924012]
12. Green MJ, Levi BH. The era of "e": the use of new technologies in advance care planning. Nurs Outlook 2012;60(6):376-383. [doi: 10.1016/j.outlook.2012.08.005] [Medline: 23141197]

13. Howard M, Robinson CA, McKenzie M, et al. Effect of an interactive website to engage patients in advance care planning in outpatient settings. Ann Fam Med 2020 Mar;18(2):110-117. [doi: 10.1370/afm.2471] [Medline: 32152014]
14. van der Smissen D, van der Heide A, Sudore RL, Rietjens JAC, Korfage IJ. Measuring advance care planning behavior in Dutch adults: translation, cultural adaptation and validation of the Advance Care Planning Engagement Survey. BMC Med Res Methodol 2021 Sep 25;21(1):194. [doi: 10.1186/s12874-021-01389-5] [Medline: 34563115]
15. Dupont C, Smets T, Monnet F, et al. Publicly available, interactive web-based tools to support advance care planning: systematic review. J Med Internet Res 2022 Apr 20;24(4):e33320. [doi: 10.2196/33320] [Medline: 35442207]
16. Monnet F, Dupont C, Smets T, et al. An advance care planning website for people with dementia and their family caregivers: protocol for a development and usability study. JMIR Preprints. Preprint posted online on Mar 2, 2023. [doi: 10.2196/preprints.46935]
17. Dupont C, Monnet F, Pivodic L, et al. Evaluating an advance care planning website for people with dementia and their caregivers: protocol for a mixed method study. Digit Health 2023;9:20552076231197021. [doi: 10.1177/20552076231197021] [Medline: 37654712]
18. Dupont C, Smets T, Monnet F, et al. A website to support people with dementia and their family caregivers in advance care planning: results of a mixed-method evaluation study. Patient Educ Couns 2024 Oct;127:108357. [doi: 10.1016/j.pec.2024.108357] [Medline: 38943816]
19. Vroegtijdige zorgplanning (VZP) & dementie [Website in Dutch]. URL: https://www.vroegtijdigezorgplanningbijdementie.be/ [accessed 2025-02-05]
20. Fahner JC, Beunders AJM, van der Heide A, et al. Interventions guiding advance care planning conversations: a systematic review. J Am Med Dir Assoc 2019 Mar;20(3):227-248. [doi: 10.1016/j.jamda.2018.09.014] [Medline: 30471947]
21. Dupont C, Smets T, Monnet F, Eneslätt M, Tishelman C, Van den Block L. The cultural adaptation of the go wish card game for use in Flanders, Belgium: a public health tool to identify and discuss end-of-life preferences. BMC Public Health 2022 Nov 17;22(1):2110. [doi: 10.1186/s12889-022-14523-9] [Medline: 36397020]
22. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008;28(5). [doi: 10.18637/jss.v028.i05]
23. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Soft 2014;61(6). [doi: 10.18637/jss.v061.i06]
24. Janssenswillen G, Depaire B, Swennen M, Jans M, Vanhoof K. bupaR: enabling reproducible business process analysis. Knowl Based Syst 2019 Jan;163:927-930. [doi: 10.1016/j.knosys.2018.10.018]
25. Potts C, Bond RR, Jordan JA, et al. Process mining to discover patterns in patient outcomes in a Psychological Therapies Service. Health Care Manag Sci 2023 Sep;26(3):461-476. [doi: 10.1007/s10729-023-09641-8] [Medline: 37191758]
26. Sellars M, Chung O, Nolte L, et al. Perspectives of people with dementia and carers on advance care planning and end-of-life care: a systematic review and thematic synthesis of qualitative studies. Palliat Med 2019 Mar;33(3):274-290. [doi: 10.1177/0269216318809571] [Medline: 30404576]
27. Núñez-Naveira L, Alonso-Búa B, de Labra C, et al. UnderstAID, an ICT platform to help informal caregivers of people with dementia: a pilot randomized controlled study. Biomed Res Int 2016;2016:5726465. [doi: 10.1155/2016/5726465] [Medline: 28116300]
28. Kajiyama B, Thompson LW, Eto-Iwase T, et al. Exploring the effectiveness of an Internet-based program for reducing caregiver distress using the iCare Stress Management e-Training Program. Aging & Mental Health 2013 Jul;17(5):544-554. [doi: 10.1080/13607863.2013.775641]
29. Ryan T, Amen KM, McKeown J. The advance care planning experiences of people with dementia, family caregivers and professionals: a synthesis of the qualitative literature. Ann Palliat Med 2017 Oct;6(4):380-389. [doi: 10.21037/apm.2017.06.15] [Medline: 28754049]
30. Piers R, Albers G, Gilissen J, et al. Advance care planning in dementia: recommendations for healthcare professionals. BMC Palliat Care 2018 Jun 21;17(1):88. [doi: 10.1186/s12904-018-0332-2] [Medline: 29933758]
31. Bastoni S, Wrede C, da Silva MC, et al. Factors influencing implementation of eHealth technologies to support informal dementia care: umbrella review. JMIR Aging 2021 Oct 8;4(4):e30841. [doi: 10.2196/30841] [Medline: 34623314]
32. Prescop KL, Dodge HH, Morycz RK, Schulz RM, Ganguli M. Elders with dementia living in the community with and without caregivers: an epidemiological study. Int Psychogeriatr 1999 Sep;11(3):235-250. [doi: 10.1017/s1041610299005803] [Medline: 10547124]

Abbreviations

ACP: advance care planning

Edited by AF Abdulai; submitted 18.05.24; peer-reviewed by C Clavien, JD Bacsu; revised version received 17.10.24; accepted 21.11.24; published 11.02.25.

Please cite as:
Dupont C, Smets T, Potts C, Monnet F, Pivodic L, De Vleminck A, Van Audenhove C, Mulvenna M, Van den Block L
Uncovering Specific Navigation Patterns by Assessing User Engagement of People With Dementia and Family Caregivers With an Advance Care Planning Website: Quantitative Analysis of Web Log Data
JMIR Aging 2025;8:e60652
URL: https://aging.jmir.org/2025/1/e60652
doi:10.2196/60652

Original Paper

The PDC30 Chatbot—Development of a Psychoeducational Resource on Dementia Caregiving Among Family Caregivers: Mixed Methods Acceptability Study

Sheung-Tak Cheng[1], PhD; Peter H F Ng[2,3], PhD

[1]Department of Health and Physical Education, The Education University of Hong Kong, Tai Po, China (Hong Kong)

[2]Department of Computing, Faculty of Computer and Mathematical Sciences, Hong Kong Polytechnic University, Hung Hom, China (Hong Kong)

[3]Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hung Hom, China (Hong Kong)

Corresponding Author:
Sheung-Tak Cheng, PhD
Department of Health and Physical Education
The Education University of Hong Kong
10 Lo Ping Road
Tai Po
China (Hong Kong)
Phone: 852 29486563
Email: takcheng@eduhk.hk

Abstract

Background: Providing ongoing support to the increasing number of caregivers as their needs change in the long-term course of dementia is a severe challenge to any health care system. Conversational artificial intelligence (AI) operating 24/7 may help to tackle this problem.

Objective: This study describes the development of a generative AI chatbot—the PDC30 Chatbot—and evaluates its acceptability in a mixed methods study.

Methods: The PDC30 Chatbot was developed using the GPT-4o large language model, with a personality agent to constrain its behavior to provide advice on dementia caregiving based on the Positive Dementia Caregiving in 30 Days Guidebook—a laypeople's resource based on a validated training manual for dementia caregivers. The PDC30 Chatbot's responses to 21 common questions were compared with those of ChatGPT and another chatbot (called Chatbot-B) as standards of reference. Chatbot-B was constructed using PDC30 Chatbot's architecture but replaced the latter's knowledge base with a collection of authoritative sources, including the World Health Organization's iSupport, By Us For Us Guides, and 185 web pages or manuals by Alzheimer's Association, National Institute on Aging, and UK Alzheimer's Society. In the next phase, to assess the acceptability of the PDC30 Chatbot, 21 family caregivers used the PDC30 Chatbot for two weeks and provided ratings and comments on its acceptability.

Results: Among the three chatbots, ChatGPT's responses tended to be repetitive and not specific enough. PDC30 Chatbot and Chatbot-B, by virtue of their design, produced highly context-sensitive advice, with the former performing slightly better when the questions conveyed significant psychological distress on the part of the caregiver. In the acceptability study, caregivers found the PDC30 Chatbot highly user-friendly, and its responses quite helpful and easy to understand. They were rather satisfied with it and would strongly recommend it to other caregivers. During the 2-week trial period, the majority used the chatbot more than once per day. Thematic analysis of their written feedback revealed three major themes: helpfulness, accessibility, and improved attitude toward AI.

Conclusions: The PDC30 Chatbot provides quality responses to caregiver questions, which are well-received by caregivers. Conversational AI is a viable approach to improve the support of caregivers.

(JMIR Aging 2025;8:e63715) doi:10.2196/63715

KEYWORDS

Alzheimer; caregiving; chatbot; conversational artificial intelligence; dementia; digital health; health care technology; psychoeducational; medical innovations; language models; mobile phone

Introduction

Background

How to leverage technological innovation to improve support in the protracted journey of caring for someone with dementia has been identified as one of the top priorities by dementia care researchers, practitioners, and policy makers [1,2]. Dementia family caregiving is widely regarded as a prototype of chronic stress [3,4]. With the increasing loss of abilities in persons with dementia, family members transition from a "companion" to a more heavy-duty caregiver. Initially, family members may want to know more about the condition and what to plan ahead of time. As the relative's condition worsens, tips about handling different behaviors or situations and helping the relative with activities of daily living become more relevant, with end-of-life issues on the horizon further down the road. Thus, caregivers' needs change over time but are often given a time-limited (eg, 8 weekly sessions) standardized training program disregarding their specific concerns [5,6]. While some open-minded caregivers may be curious to find out more about the subject matter, others feel compelled to go through materials irrelevant to their immediate concerns (or things they already know) [7]. Moreover, traditional interventions are packaged to be delivered in regular blocks of time, which may not fit into the busy and often unpredictable schedule of many caregivers [5,6]. In addition, traditional intervention modalities often rely on face-to-face contact, a problem brutally exposed during the COVID-19 pandemic [8,9], but a problem that is generally applicable to those in rural areas.

Although over two-thirds of caregivers in a large UK survey [10] indicated an interest in using technological tools to support their health and well-being, as well as their caregiving role, few such tools are available for dementia family caregivers. A major reason cited by the respondents in support of technological support was "being able to use digital tools quickly to find answers on a regular basis" [10]. Relatedly, this study explores how caregivers react to a chatbot based on generative artificial intelligence (AI) technology, which provides information and advice on demand and can be accessed anytime anywhere.

AI Applications for Dementia Care

AI (especially natural language processing) refers to a collection of algorithms that gets computers or machines to perform autonomous actions that require human intelligence (eg, understanding complex semantics and intent, reasoning, decision-making, and self-correction). In the field of health care, the advent of AI has led to a wide range of applications, such as behavioral monitoring, risk prediction, screening, triage, diagnosis, rehabilitation, health advice chatbot, robotics, and program planning [11-21]. Such applications may be used to assist or replace human activity, or even to improve performance by enhancing efficiency and reducing human error [22-26]. Thus, AI has the potential to fill some service gaps for families with dementia as societies grapple with this public health crisis.

That said, AI is not without drawbacks. For end users, the greatest concerns are the accuracy of the output and the quality of decision-making [27,28]. An AI program is only as good as the algorithm and the data used to train it. For example, a health risk prediction program in the United States classified Black people as having lower risk profiles, compared with White people with comparable health status, resulting in many Black people being denied the proper service. The reason was due to the algorithm using health care cost as the proxy for how sick the person was, without taking into account the fact that Black people have historically been an underserved population [29].

Two reviews provide valuable insights into the development and use of AI chatbots and technology-driven solutions to support dementia care. Hoel et al [30] explored various technology-driven solutions such as social robots and tablet applications that provide activities or interactions to engage persons with dementia, most of whom live in residential care settings. Although the primary targets were persons with dementia, interactions with family caregivers were found to be enhanced through the involvement of the caregiver in the activity (eg, reminiscence therapy). Through more enjoyable interaction, the level of stress felt by family caregivers might be reduced.

The systematic review by Ruggiano et al [9] focused on chatbots, conversational AI agents that use natural language processing and large language models (both machine learning algorithms) to simulate complex human conversation [31,32]. They found 6 commercially available chatbots designed to assist people with dementia and their caregivers. Focusing on the functions and quality of these tools, their review highlighted the potential of chatbots to offer educational content for caregivers (and memory aids for patients) through accessible platforms like mobile apps and voice-activated devices. These chatbots aimed to provide users with timely information and interactive features to engage people with dementia, but a common drawback was the lack of peer-reviewed, evidence-based educational material, which undermines their trustworthiness. As a matter of fact, none of the chatbots were deployed following empirical user evaluation [9]. The information provided was often not rooted in rigorous scientific research, leaving caregivers unsure about the accuracy and reliability of the advice they receive. Based on limited programmed content, these chatbots inevitably constrain interaction with users and provide advice on a narrow range of questions [9].

Another issue is accessibility. Many chatbots require specific phrases or commands to operate effectively, limiting their flexibility and usability for a diverse range of caregivers. This rigidity can be frustrating, especially for those who are not technologically proficient. Additionally, most specialized chatbots provide only standard answers to preset questions and cannot analyze user intent and emotion. They rely heavily on users to provide specific contextual information for continued engagement.

This Study

In view of the limitations of existing chatbots, this study reports on a newly built chatbot, called PDC30 Chatbot, which serves as a care adviser on dementia caregiving using generative AI technology and an evidence-based, comprehensive knowledge base. By simply providing information and advice, without involvement in decision-making, it avoids potential risks of AI applications mentioned above. At the same time, the use of a

validated knowledge base enhances the quality of the answers and trust among users, with its information extracted by generative AI to produce a new and original response to each caregiver enquiry. As such, this new chatbot offers greater flexibility in input handling, allowing caregivers to engage in natural, conversational interactions without the need for rigid commands. Furthermore, through remarkably improved emotional and cognitive support by virtue of generative AI technology, the chatbot is capable of delivering empathetic responses and practical caregiving strategies without requiring constant caregiver input. This makes it a more autonomous and complete tool for addressing the emotional, cognitive, and practical demands of dementia caregiving.

In the following, we describe in detail the construction of the PDC30 Chatbot and its acceptability from a user perspective. A new chatbot would have little value if there is no evident advantage over other resources. Thus, the chatbot's performance in relation to handling common questions concerning dementia and caregiving was examined by comparing its responses to those of two other chatbots including the popular ChatGPT. Upon establishing its performance value, the PDC30 Chatbot was subject to an assessment of acceptability after a 2-week use by dementia family caregivers.

Methods

Chatbot Development

PDC30 Chatbot

We use the latest GPT-4o large language model by OpenAI [33], with the following prompt to define the personality traits of the chatbot (ie, a personality agent in large language models): You are a professional counselor providing advice on caregivers' emotional issues and caregiving challenges. The knowledge base is the Positive Dementia Caregiving in 30 Days Guidebook, an abridged and updated version of the Benefit-Finding Intervention manual (a psychoeducation program with an emphasis on searching for positive meanings in caregiving) which has been found to reduce caregiver burden and depression, with moderate to large effect sizes, up to 12 months [34-39] (the Benefit-Finding Intervention is a workshop-based program with a lengthy instruction manual written for the trainers. The manual was rewritten in simple language and a more concise form for general public consumption and became the knowledge base for this chatbot). After providing an answer, the chatbot is programmed to reference the source of the ideas, so as to encourage users to do more in-depth reading.

In other words, the chatbot would not answer any question unrelated to dementia and caregiving. As a generative AI chatbot, it can formulate answers to a wide range of questions based on these knowledge bases, rather than simply providing preset answers to selected questions. When irrelevant questions are asked, it would say "Sorry, I cannot answer your question." The chatbot, accessible on any device with internet access (including phone, tablet, or computer), was built on botpress.com, an open-source platform for conversational AI solutions. The chatbot serves as one of the components of a self-guided, automated web-based intervention program called Positive Dementia Caregiving in 30 Days (PDC30), which is undergoing evaluation in a global randomized controlled trial (ClinicalTrials.gov identifier NCT06409455). Hence, we call it the PDC30 Chatbot. Currently, this web-based intervention program, the chatbot included, works only in the English language.

ChatGPT and Chatbot-B as Comparators

To evaluate how well the PDC30 Chatbot works, the same questions (see below) were fed into ChatGPT-4o (as a standard of reference) [33] and another self-constructed chatbot (called Chatbot-B for convenience). As ChatGPT is well-known, we focus on introducing Chatbot-B.

Chatbot-B's design is identical to that of the PDC30 Chatbot, but the reference materials are different. The knowledge base for Chatbot-B consists of (1) 154 caregiving-related web pages by the Alzheimer's Association, (2) 17 dementia caregiving-related web pages by the US National Institute on Aging, (3) 14 caregiving topic-based manuals by the Alzheimer's Society, United Kingdom, (4) World Health Organization's iSupport manual [40], and (5) By Us For Us Guides, a collection of 15 Canada-based documents written by people with dementia and their caregivers [41] (if a document was not optimized for automated processing, its content was manually extracted and saved as a plain text file to improve the chatbot's ability to interpret and analyze the information). In other words, Chatbot-B represents a rather comprehensive and authoritative knowledge base from which answers are drawn. In this sense, Chatbot-B is like an assistant surfing the internet for relevant materials on behalf of the user and summarizes the main points for the user in far less time than the user surfing the internet himself or herself.

Together, ChatGPT and Chatbot-B provide strong reference points for assessing the quality of PDC30 responses. ChatGPT is an existing tool that can be used by caregivers in countries where it is available. It is trained on an enormous text retrieved from the internet but is known to have the risk of providing false information when such information exists on the internet. This reliance on internet-based sources makes it vulnerable to inaccuracies, particularly when it draws from unreliable content. Additionally, the chatbot may struggle with context or nuance, offering responses that are overly general or missing critical domain-specific details, which can be problematic in sensitive caregiving scenarios [9]. Chatbot-B, on the other hand, is a new chatbot dedicated to dementia and caregiving topics using authoritative materials. A caveat needs to be mentioned. There are pros and cons of including many reference materials. Up to a certain point, the benefit of including more texts levels off given the redundancy across their content. Processing time may be lengthened as the knowledge base expands in size. Different texts may also offer advice that contradicts each other. Hence, a chatbot based on a single text (ie, the PDC30 Guidebook) may not necessarily fare worse.

Testing Materials and Chatbot Responses

We constructed 21 questions commonly asked by family members by surfing the internet and subjected the three chatbots to the same questions to see how their answers compare. The

responses provided by PDC30 Chatbot and Chatbot-B are reproduced in Multimedia Appendix 1, whereas those by ChatGPT are shown in Multimedia Appendix 2. Note that due to the use of generative AI, the answers provided to the same question will vary slightly from time to time; thus, these are not fixed answers, but rather one sample of a range of possible answers by the respective chatbot. Moreover, it should be mentioned that the boldfaced headings in front of the bullet points were created by the chatbots; they did not exist in the original sources. Interestingly, at one point, the PDC30 Chatbot addressed the first author ("Dear Tak," response to the question "Why is my wife acting like a different person") and that was the only instance any of the chatbots addressed the user by name when answering questions.

For PDC30 Chatbot and Chatbot-B, one can see that despite using proper nouns (eg, mom and husband) and pronouns (eg, her), the chatbots had no problem understanding they were the care recipients. The wording in some questions was intentionally nonspecific (to mimic everyday conversation) and GPT-4o, when constrained to answer questions related to dementia and caregiving, had no difficulty grasping the meaning in the context of dementia caregiving. On the contrary, ChatGPT expectedly needed information on the specific context to construct more relevant responses (for instance, when asked about preparation for the future, it talked about financial investment unless being told that the question concerned a family having a relative newly diagnosed with dementia). Hence questions for ChatGPT were elaborated accordingly. By comparison, questions posed to PDC30 Chatbot and Chatbot-B mimicked natural conversations a lot more.

In terms of answers, occasionally there are bullet points, the relevance of which to the question is not immediately apparent (eg, Chatbot-B's points "utilize online services" and "stay engaged and active" to the question "What preparation should we make for the future?"). This issue appeared to be more common for ChatGPT, including, but not limited to: considering home safety and daily activities for the preparation for future question; encouraging activity involvement for the question on communication skills; discussing financial or legal planning and communication skills for the question on managing personal feelings and stress; asking the caregiver to join support groups, find respite, consider environmental safety, and to take breaks for himself or herself when the question was how to deal with mom's apathy; and getting adult day care, power of attorney, and medical directives when being asked how to allow husband to wander safely. Such answers may well confuse caregivers, limiting the value of ChatGPT as a consultation resource.

Moreover, ChatGPT responses tended to be repetitive, while missing some key advice to caregivers, such as avoiding confrontation and giving due recognition to the disease as the real causal agent for problematic behaviors. On the whole, the responses by PDC30 Chatbot and Chatbot-B were more concise, specific, and to the point. Considering the fact that these chatbots were providing advice on complicated matters, we think their overall performance was quite good—the results support chatbots using generative AI technology as a viable approach to offer advice to caregivers, and the results are better with topic-focused chatbots. For this reason, we focus on the relative performance of PDC30 Chatbot and Chatbot-B below.

The answers provided by these two chatbots were surprisingly similar. In fact, the answers to the question "I am mad with myself; I made so many mistakes" were identical between the two chatbots, though PDC30's answer was probably clearer with more elaboration. Though there is no straightforward way to ascertain the relative quality of the answers, we are of the opinion that compared with PDC30 Chatbot, Chatbot-B's performance was inferior in relation to the questions "I cannot accept the idea that my Mom has dementia and may leave me one day" and "I am so frustrated I can't control my emotions; I would even take out my irritation on her." One difference between the two chatbots does stand out, which is that when it comes to questions related to emotional or mental health issues (technically the above question about not being able to accept mom's condition is also an emotionally laden question), the PDC30 Chatbot provided a greater variety of suggestions. For example, pleasant event scheduling and alternative thinking are important coping strategies but were mentioned by the PDC30 Chatbot only. Other strategies including positive meaning-making and tackling unrealistic assumptions about the care recipient were also emphasized a lot more by the PDC30 Chatbot. These differences are notable, as psychoeducational interventions that include tactics taken from psychotherapy have been found, in a comprehensive meta-analysis of 131 randomized controlled trials, to be much more beneficial for dementia caregivers (especially in terms of relieving depression and promoting self-efficacy and positive gains) than psychoeducation without such tactics [2]. On the whole, the results support the value of the PDC30 Chatbot as a "counselor" for dementia caregivers. Typically, an answer was provided in 20 seconds or less, despite the volume to be scanned, suggesting that a generative AI chatbot is a rather convenient way for caregivers to obtain basic information and advice.

Some brief remarks about Chatbot-B are warranted before moving on to the main acceptability study. In terms of referencing, Alzheimer's Association materials were used in formulating responses to 14 of the 21 questions; Alzheimer's Society and By Us For Us documents were each used for 12 questions; National Institute on Aging web pages, 9 questions; and iSupport, 8 questions. Thus GPT-4o did scan the entire knowledge input before drafting its answers. Note also that Chatbot-B works only if the URLs are up to date. This can make it a less preferable choice as an intervention tool, compared with the PDC Chatbot which has a stable and well-defined knowledge source.

Acceptability Study for PDC30 Chatbot—Design and Procedure

In light of the positive results for the PDC30 Chatbot, we proceeded to the next phase to evaluate the acceptability of the chatbot. A mixed methods study was conducted, in which caregivers were shown the PDC30 Chatbot's responses to the 21 questions above, and asked to use the chatbot (accessible using a hyperlink) for two weeks, at least once a day (as this study was focused on the evaluation of the chatbot, the other components of the PDC30 intervention were not included). At

the end of 2 weeks, participants indicated, on a self-report questionnaire, their frequency of using the chatbot on a scale of 1=almost never, 2=several times a week but less than daily, 3=once a day, and 4=more than once a day. They rated the chatbot using the following questions: (1) the chatbot is easy to use, (2) the answers are easy to understand, (3) the answers are helpful, (4) overall, you are satisfied with the chatbot, and (5) you would recommend the chatbot to other caregivers (all rated from 1=strongly disagree to 5=strongly agree). Participants were also asked to write down any other thoughts they had about the chatbot; this written feedback was subject to thematic analysis [42].

Participants

Participants were recruited through posting notices on campus. The inclusion criteria are (1) aged 18 years or older, (2) self-identification as providing care to a relative with dementia, and (3) self-reported fluency in English. There was no exclusion criterion. All participants were relatives of students and staff who were taking care of a family member with dementia and who provided informed consent to participate. We stopped recruitment after the 21st participant because data saturation had evidently been reached, as the same themes kept repeating. Participants were not financially compensated but were given continuing access to the PDC30 Chatbot after the study.

Ethical Considerations

The study was approved by the Human Research Ethics Committee of The Education University of Hong Kong (reference 2021-2022-0077). Written informed consent was obtained from all participants.

Data Analysis

The two authors independently read the participants' written feedback to first become familiar with the entire set of qualitative data. After highlighting key points and making notes, they generated initial codes for each written feedback. Themes were then extracted by identifying similar codes and patterns using an inductive approach. The two authors compared their work and found that similar themes were identified. The themes were finalized and articulated [42]. As for the Likert-type questions, data were summarized using descriptive statistics.

Results

Overview

Participant characteristics and their ratings of the chatbot can be found in Table 1. Understandably, participants were predominantly child caregivers (10 daughters and 4 sons). Two caregivers were the daughters-in-law of the care recipient. Only four were spouses (one wife and three husbands) and one was a sister.

As can be seen from Table 1, the responses were overwhelmingly positive. A total of 15 out of 21 (71%) participants used the chatbot more than once per day (overall mean 3.62, SD 0.67). Participants thought the chatbot was very easy to use and most gave the highest rating (mean 4.52, SD 0.68). They also thought the advice provided was quite helpful (mean 4.29, SD 0.56) and relatively easy to understand (mean 3.81, SD 0.98). Given such positive experiences, it is no wonder that they felt rather satisfied with the chatbot (mean 4.05, SD 0.59). A total of 16 out of 21 (76%) participants would recommend or strongly recommend it to other caregivers (mean 4.24, SD 0.83). These ratings were more or less substantiated by the participants' written qualitative feedback (Table 2). Thematic analysis of the written comments revealed 3 recurring themes.

Table 1. Acceptability study sample and their ratings of the PDC30 Chatbot.

Caregiver	Age	Sex	Relationship	Q1[a]	Q2[b]	Q3[c]	Q4[d]	Q5[e]	Q6[f,g]
A	40	F	Child	4	5	4	5	4	5
B	58	M	Child	4	4	5	4	4	5
C	51	F	Child	4	5	5	5	5	4
D	36	F	Child	3	5	5	4	4	5
E	73	F	Spouse	2	5	2	3	3	3
F	50	F	Child	4	5	5	4	5	5
G	70	M	Spouse	4	5	3	4	5	4
H	46	M	Child	4	4	4	4	4	5
I	42	F	Child	4	5	4	5	4	5
J	50	F	Child	4	4	3	5	4	5
K	43	F	Child-in-law	4	5	5	4	3	4
L	44	F	Child	3	4	3	4	4	4
M	48	F	Child	2	5	4	5	3	3
N	55	F	Child-in-law	4	5	5	4	4	5
O	50	F	Child	3	4	4	4	5	5
P	72	M	Spouse	4	4	3	4	4	5
Q	71	M	Spouse	4	3	2	4	4	4
R	54	M	Child	4	5	4	5	4	3
S	68	F	Sibling	4	3	3	5	4	3
T	50	M	Child	4	5	4	4	4	3
U	55	F	Child	3	5	3	4	4	4
Mean (SD)	53.62 (11.18)	—[h]	—	3.62 (0.67)	4.52 (0.68)	3.81 (0.98)	4.29 (0.56)	4.05 (0.59)	4.24 (0.83)

[a]Q1: frequency of use.

[b]Q2: whether easy to use.

[c]Q3: answers easy to understand.

[d]Q4: answers helpful.

[e]Q5: satisfaction.

[f]Q6: likelihood to recommend to others.

[g]Except for Q1 which is scored 1-4, all other questions are scored 1-5, with higher scores indicating a better experience.

[h]Not applicable.

Table 2. Participants' written comments (reproduced in full) about the PDC30 Chatbot.

Caregiver	Written comments
A	I love it! I have to work and take care of my mom. My sister do not help. I feel very lonely and not sure what to do at times. Now I can have this chatbot to talk to any time I want - it feels like a friend. And it gives me ideas about how to handle my mom too.
B	Don't feel I should go get help from the centre's social worker again. The chatbot's advise is better! And it works 24 hours a day! I actually feel so stupid now that I never trusted AI. I am really surprised how good it is.
C	The answer refer to a guidebook but we don't have it. That part is a bit confusing. That said, I really like the fact that I can now ask anything any time.
D	Initially I have reservation. I have tried many chatbots (like banks) and they are so annoying because they pretended to help but only gave you standard replies not addressing your query. This one is different. Impressed!
E	I am very busy because my husband fell. So I don't use it much. I appreciate the advise but some suggestions are not easy to follow. But I find it very easy to use. My first thought it is strange talking to a robot. But after using it, I don't feel that anymore. When ok, I will use it more.
F	Can I keep using the chatbot, please? When my nephew mentioned this chatbot to me, I thought it is a joke. Talking to a machine? But I use it every day now. Please let me use it.
G	How come not available earlier? I wish had it 2 years ago when I was most helpless!
H	Did Tom Cruise's boss said drones do not sleep but pilots do? Well, this is a very good drone.
I	This is the first time a chatbot is REALLY answering my questions! There is some wait but waiting time reasonable.
J	Professor, thanks indeed for giving me a chance to use this chatbot. It is very helpful!
K	Most answers are relevant and informative. Quick too. Better than I thought.
L	It say it is not a therapist but it is a therapist. It help me calm down when I was very very frustrated! Thanks. I will try to use it more.
M	A good companion. Answer quality is good. But I use it only when I could not find someone to talk to. It's a backup when no one is available.
N	I can screen cap the answer and send to my husband. Then we discuss how to handle my mother-in-law. I feel more confident because I have both knowledgeable advice and husband support.
O	Frankly, I was very skeptical in the beginning. But I gradually like it more. I have a smart speaker but it does not give good answer to this type of questions.
P	Never used a chatbot. Very interesting. Good information. My wife doesn't know what I do on my phone when I am looking for help. If I call my daughter, she would hear what we said and feel upset.
Q	Some terms are difficult. I need to check dictionary. But very convenient. 24 hour service. My wife has frequent difficult behavior and I like the suggestion to ignore or take a break. Some tips for bathing and serving meal helped.
R	It's a good chatbot. I love it but 2 weeks too short to say how good it really is.
S	Not used to talk to machine. Often I have a problem but forget the chatbot is there. Forget can use my phone. Then afterwards when I see my phone, I remember! Then I ask question and get some helpful answer but sometime too late.
T	I work in IT field and so this is no surprise to me. I am glad something like this is finally available because caregiving is hard! I did not get any nonresponse and response time is very good. The answers are specific with enough elaboration and minimal repetition. These are what determine good user experience.
U	Amazing. It answer all my questions so far. But my father is mild stage. May be my questions not hard enough. Interesting to see if it can help me when things get more tough.

Theme 1: Helpfulness

This is the most dominant theme which was explicitly (caregivers A, B, I, J, L, K, M, N, P, Q, and U) or implicitly (caregivers D, F, G, H, O, R, S, and T) mentioned by 19 (90%) participants. Four (caregivers D, I, O, and T) said the chatbot's answers were appropriate, informative, and specific, unlike other chatbots that provided inadequate, irrelevant, or standardized replies. Caregiver L specifically mentioned that the chatbot helped her to calm down when severely frustrated. Her comments were echoed by caregiver Q who also referred to helpful practical suggestions by the chatbot. Caregiver F, despite being informed at the study's outset, forgot that the chatbot would continue to be available, and asked if she could keep on using it after the testing period, a gesture suggestive of the helpfulness of the chatbot. Another participant (caregiver G), who cared for his wife, lamented the late coming of this resource and imagined what difference it would have made to his situation if it had been available 2 years earlier, again testifying to the perceived helpfulness of the chatbot. By comparing the chatbot to a "very good drone," caregiver H might be saying that the chatbot got the job done (ie, providing good advice), a point verified in a subsequent email conversation with the caregiver. Perhaps a bit exaggerating, caregiver B even complimented the chatbot by saying that its advice was better than the social worker's. One wife caregiver (E), however, thought that some of the suggestions by the chatbot were not easy to follow, which was not entirely surprising as behavioral

and cognitive change takes time and participants did not have the full intervention program to assist them. On the positive side, this comment suggests that the answers were relevant, just not easy to follow through.

Theme 2: Accessibility

In total, 11 (52%) caregivers mentioned, in one way or another, the usability of the chatbot, including being user-friendly (caregiver E), lack of (or low rates of) nonresponse (caregiver T), reasonable waiting time before getting a response (caregivers I, K, and T), and most of all, the chatbot's 24/7 accessibility (caregivers A, B, C, H, and Q). Caregiver H even paraphrased a conversation in the movie "Top Gun: Maverick" to illustrate this point.

Unexpectedly, two caregivers mentioned functional merits unrelated to the chatbot itself. Caregiver N talked about the convenience of saving the chatbot's responses and sharing them with other family members. This action, as alluded to by this caregiver, might encourage more involvement by other family members and support to the main caregiver. In addition, caregiver P made an interesting point about the privacy afforded by using the chatbot on his smartphone. His comment was a reminder of the dilemma faced by many caregivers, especially those in crowded living conditions such as the case in Hong Kong, when trying to seek help when the care recipient is nearby. Conversations, whether over the phone or face-to-face, may be overheard by the care recipient, who, in turn, reacts with more behavior problems. The chatbot operated on the phone offers caregivers private space to get help when necessary.

Theme 3: Attitude Toward AI

Five participants (caregivers B, D, E, F, and O) expressed that they were initially skeptical of AI but the experience changed their perception of it. For example, caregiver B started with a mistrust of AI while caregiver F initially found seeking advice from a machine a ridiculous idea. After trying it, both became excited by the technology and were using the chatbot more than once a day. A related sentiment was expressed by another caregiver, S. She was not used to talking to a machine, and after having access to the chatbot, kept forgetting about it. Yet, the undertone of the feedback was that she found the chatbot responses to be useful and lamented that occasionally she had not taken timely advantage of the resource.

Not everyone was as receptive to the technology. Caregiver M did not reject the chatbot but reserved it for occasions when she could not find people to talk to. Nevertheless, the feedback overall suggested that experience with the chatbot induced a favorable attitude toward it in those who questioned its value to begin with.

Discussion

Principal Findings

This study demonstrates that using a chatbot to help dementia caregivers is a viable approach. Taking advantage of AI development, we were able to build the PDC30 Chatbot that functions as a "counselor" to caregivers (ie, its purpose), by applying a personality agent to constrain the chatbot's behavior and by feeding it with an appropriate knowledge base. The generative AI algorithm summarized the points using "its own" words and organized them using headings to facilitate reading. At first, it might seem that feeding it with more resource materials would help it formulate better responses—for this reason, two other chatbots were included for comparison, namely ChatGPT and Chatbot-B incorporating insider perspectives (those from patients and caregivers) and guidelines by several authoritative agencies. In our testing using 21 common questions by caregivers, there was indeed a tendency for Chatbot-B to generate slightly longer responses than PDC30 Chatbot, though this was not always the case. Nevertheless, length per se does not determine quality, as longer answers may contain irrelevant and repetitive points, as we have seen, especially in the case of ChatGPT. On the whole, we think that PDC30 Chatbot and Chatbot-B were superior to ChatGPT and performed similarly, with PDC30 having a further edge on emotional and mental health issues. In addition, the PDC30 Chatbot was favorably received in an acceptability study of 21 dementia family caregivers after using it more than once a day for two weeks. The chatbot was considered user-friendly, with its responses helpful and easy to understand. Their written feedback about the chatbot centered on three themes: it was helpful and accessible, and it improved their perception of AI as a helping agent. Overall, the experience was rather satisfactory and the participants would strongly recommend it to other caregivers.

Both PDC30 Chatbot and Chatbot-B provide citations of the sources they refer to when formulating responses. A caveat needs to be mentioned. While the citation system may prompt caregivers to do more reading of the relevant materials, it is good only to the extent that the sources can be readily located. As explained in Table 1 footnote, GPT-4o cannot name the author of the web pages cited. If the source is a document, this was handled by the way the document is named, such as "Alzheimer's Society Making Your Home Dementia Friendly"—in other words, the citation is to the name of the file used. If it is just a web page, GPT-4o provides a citation using the first title that appears on the web page; the agency or author producing the web page is omitted—this is primarily due to limitations in GPT-4o's design and functionality. First, GPT-4o cannot read logos or metadata from images on web pages, where authors' names and relevant citation details are often embedded. If the website does not contain alternative text, GPT-4o cannot read the images. Second, while the model processes textual data, it may only effectively distinguish between content and metadata, such as author names or publication details, if explicitly mentioned in the text. The website needs to state the metadata clearly. Additionally, the training data (ie, web page metadata) may not include structured citation formats, making it difficult for the model to recognize and extract authors' names reliably. Furthermore, the design of GPT-4o may include privacy considerations that prevent it from naming individuals unless explicitly included in the training data or provided in a clear, structured format. To address this issue, converting all relevant web pages into text files with clear filenames and structured citation information can help ensure that author names and other citation details are correctly identified and referenced by the model. On the contrary, the PDC30 Chatbot is to be used

together with the Guidebook in actual operation—both are to be hosted on the same intervention website. Thus, the citation problem mentioned above does not apply.

Perhaps a more important consideration in terms of operation is that Chatbot-B requires constant monitoring and updating, should organizations add new web pages, remove certain web pages, or change the URLs of the web pages—there are indeed a very large number of web pages to monitor. Worse, changes to the content of a web page with the same URL will be hard to detect, unless a "last updated date" is provided. By comparison, the PDC30 Chatbot is much easier to maintain, as it is based on only one text.

The PDC30 Chatbot was built with the intention to be included in a randomized controlled trial, and so we subject it to an evaluation of acceptability by caregivers. All but one caregiver provided highly positive feedback about the chatbot; the only exception came from a wife caregiver whose use of the chatbot was disrupted by her husband's fall early in the trial period (had she contacted the research team then, we would have adjusted her trial period so that she would have been able to try the chatbot at a more convenient time. Unfortunately, we were not aware of the issue until she turned in her responses). On the whole, the caregivers found the chatbot easy to use (including the wife caregiver mentioned above) and the answers were informative, relevant to their needs, and easy to understand. Many particularly appreciated the chatbot's helpfulness and 24/7 accessibility, with some further commenting on its lack of nonresponse and reasonable waiting times before responses were generated. As a result, some caregivers acquired a much more favorable attitude toward AI technology because of using the chatbot. The great majority reported high levels of satisfaction with the chatbot and would recommend it to other caregivers. Thus, the PDC30 Chatbot was well-received by caregivers who used it. The chatbot's frequent use was also a testimony to its reception among the caregivers.

Limitations

Despite the encouraging findings, a few limitations need to be mentioned. First, the sample was predominantly adult children. We do not think the paucity of spousal or sibling caregivers was due to the relative lack of digital literacy among older caregivers as there is no skill required to use the chatbot (other than typing questions into the input space). The older, spousal or sibling caregivers who participated generally found the chatbot easy to use. Rather, we think that this reflects cohort differences in the preference for digital material [43]. Furthermore, the composition of the sample is not as one-sided as it seems, as about two-thirds of the caregivers in this community are children [44]. That said, there is no doubt a certain degree of self-selection as not everyone prefers digital resources (we did have two caregivers in the sample who were initially doubtful of the technology). A future study focusing on older caregivers, especially those with little digital literacy, is warranted.

Second, the chatbot was constructed to speak English only, in alignment with the language used in the parent intervention program being evaluated in a global trial. It is probably too harsh to consider this a limitation but in the context of this study, fluent English-speaking persons do not represent the caregiver population in Hong Kong. Due to the time pressure to launch the clinical trial, we also did not recruit caregivers from English-speaking countries. Although the generalizability of the findings may be limited, the PDC30 Chatbot's responses shown in Table 1 suggest that the chatbot should be quite suitable for caregivers in other countries as well.

Conclusions

To the best of our knowledge, this is the first study reporting the development of a psychoeducational chatbot for dementia caregivers and testing its functioning and acceptability among caregivers. The convenience of delivering advice to caregivers by AI-driven chatbots is an approach that needs to be explored further by the field. Services for families with dementia are lacking, especially in resource-poor countries, and even when services are available, many caregivers do not use the services due to lack of time, services not meeting their needs, or simply wanting to locate information themselves [45]. Good chatbots may well fill some of the gaps and get help to these caregivers around the world. There is one more important feature that makes chatbots an advantageous option—only small amounts of data are transmitted each time and hence chatbots are especially suitable for low-income countries and rural areas in general where internet coverage remains an issue.

Acknowledgments

The study was supported by the General Research Fund (18622322) of the Research Grants Council of Hong Kong to the first author. The funder and the sponsor play no roles in conceptual formulation, research design and execution, data acquisition and analysis, and report write-up.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

STC conceptualized the study and designed the study and the two chatbots. STC and PHFN cosupervised the programmer to set up the chatbots. Both authors contributed to the writing of this manuscript, with STC taking the leadership. Both authors approved the submission and take responsibility for the accuracy and integrity of the work. No professional writer was employed.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Responses provided by PDC30 (Positive Dementia Caregiving in 30 Days) Chatbot and Chatbot-B.
[DOCX File , 55 KB - aging_v8i1e63715_app1.docx]

Multimedia Appendix 2
Responses provided by ChatGPT.
[DOCX File , 57 KB - aging_v8i1e63715_app2.docx]

References

1. Harvath TA, Mongoven JM, Bidwell JT, Cothran FA, Sexson KE, Mason DJ, et al. Research priorities in family caregiving: process and outcomes of a conference on family-centered care across the trajectory of serious illness. Gerontologist 2020;60(Suppl 1):S5-S13 [FREE Full text] [doi: 10.1093/geront/gnz138] [Medline: 32057081]
2. Cheng ST, Li KK, Losada A, Zhang F, Au A, Thompson LW, et al. The effectiveness of nonpharmacological interventions for informal dementia caregivers: an updated systematic review and meta-analysis. Psychol Aging 2020;35(1):55-77. [doi: 10.1037/pag0000401] [Medline: 31985249]
3. Cheng ST. Dementia caregiver burden: a research update and critical analysis. Curr Psychiatry Rep 2017;19(9):64 [FREE Full text] [doi: 10.1007/s11920-017-0818-2] [Medline: 28795386]
4. Vitaliano PP, Young HM, Zhang J. Is caregiving a risk factor for illness? Curr Dir Psychol Sci 2004;13(1):13-16. [doi: 10.1111/j.0963-7214.2004.01301004.x]
5. Cheng ST. Providing Optimal Support to Caregivers. Germany: Springer Science and Business Media LLC; 2021.
6. Cheng ST, Li K, Or PPL, Losada A. Do caregiver interventions improve outcomes in relatives with dementia and mild cognitive impairment? A comprehensive systematic review and meta-analysis. Psychol Aging 2022;37(8):929-953. [doi: 10.1037/pag0000696] [Medline: 35771500]
7. Yu Y, Xiao L, Ullah S, Meyer C, Wang J, Pot AM, et al. The experiences of informal caregivers of people with dementia in web-based psychoeducation programs: systematic review and metasynthesis. JMIR Aging 2023;6:e47152 [FREE Full text] [doi: 10.2196/47152] [Medline: 37247218]
8. Liang J, Aranda MP. The use of telehealth among people living with dementia-caregiver dyads during the COVID-19 pandemic: scoping review. J Med Internet Res 2023;25:e45045 [FREE Full text] [doi: 10.2196/45045] [Medline: 37227755]
9. Ruggiano N, Brown EL, Roberts L, Suarez VF, Luo Y, Hao Z, et al. Chatbots to support people with dementia and their caregivers: systematic review of functions and quality. J Med Internet Res 2021;23(6):e25006 [FREE Full text] [doi: 10.2196/25006] [Medline: 34081019]
10. Egan KJ, Clark P, Deen Z, Dutu CP, Wilson G, McCann L, et al. Understanding current needs and future expectations of informal caregivers for technology to support health and well-being: national survey study. JMIR Aging 2022;5(1):e15413 [FREE Full text] [doi: 10.2196/15413] [Medline: 35084339]
11. Ng PHF, Chen PQ, Sin ZPT, Lai SHS, Cheng ASK. Smart work injury management (SWIM) system: a machine learning approach for the prediction of sick leave and rehabilitation plan. Bioengineering (Basel) 2023;10(2):172 [FREE Full text] [doi: 10.3390/bioengineering10020172] [Medline: 36829666]
12. Cheng ASK, Ng PHF, Sin ZPT, Lai SHS, Law SW. Smart work injury management (SWIM) system: artificial intelligence in work disability management. J Occup Rehabil 2020;30(3):354-361. [doi: 10.1007/s10926-020-09886-y] [Medline: 32236811]
13. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digital Med 2020;3:17 [FREE Full text] [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
14. Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JMC. Clinical decision support systems for triage in the emergency department using intelligent systems: a review. Artif Intell Med 2020;102:101762. [doi: 10.1016/j.artmed.2019.101762] [Medline: 31980099]
15. Bu X, Ng PH, Xu W, Cheng Q, Chen PQ, Cheng AS, et al. The effectiveness of virtual reality-based interventions in rehabilitation management of breast cancer survivors: systematic review and meta-analysis. JMIR Serious Games 2022;10(1):e31395 [FREE Full text] [doi: 10.2196/31395] [Medline: 35225817]
16. Bu X, Ng PHF, Tong Y, Chen PQ, Fan R, Tang Q, et al. A mobile-based virtual reality speech rehabilitation app for patients with aphasia after stroke: development and pilot usability study. JMIR Serious Games 2022;10(2):e30196 [FREE Full text] [doi: 10.2196/30196] [Medline: 35389349]
17. Hird N, Osaki T, Ghosh S, Palaniappan SK, Maeda K. Enabling personalization for digital cognitive stimulation to support communication with people with dementia: pilot intervention study as a prelude to AI development. JMIR Form Res 2024;8:e51732 [FREE Full text] [doi: 10.2196/51732] [Medline: 38227357]

18. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021;4(4):e30353 [FREE Full text] [doi: 10.2196/30353] [Medline: 34874886]
19. Lee H, Chung MA, Kim H, Nam EW. The effect of cognitive function health care using artificial intelligence robots for older adults: systematic review and meta-analysis. JMIR Aging 2022;5(2):e38896 [FREE Full text] [doi: 10.2196/38896] [Medline: 35672268]
20. Wilmink G, Dupey K, Alkire S, Grote J, Zobel G, Fillit HM, et al. Artificial intelligence-powered digital health platform and wearable devices improve outcomes for older adults in assisted living communities: pilot intervention study. JMIR Aging 2020;3(2):e19554 [FREE Full text] [doi: 10.2196/19554] [Medline: 32723711]
21. Wong KLY, Hung L, Wong J, Park J, Alfares H, Zhao Y, et al. Adoption of artificial intelligence-enabled robots in long-term care homes by health care providers: scoping review. JMIR Aging 2024;7:e55257 [FREE Full text] [doi: 10.2196/55257] [Medline: 39190455]
22. Kilsdonk E, Peute LW, Riezebos RJ, Kremer LC, Jaspers MWM. Uncovering healthcare practitioners' information processing using the think-aloud method: from paper-based guideline to clinical decision support system. Int J Med Inform 2016;86:10-19 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.11.011] [Medline: 26725690]
23. Dowding D, Mitchell N, Randell R, Foster R, Lattimer V, Thompson C. Nurses' use of computerised clinical decision support systems: a case site analysis. J Clin Nurs 2009;18(8):1159-1167. [doi: 10.1111/j.1365-2702.2008.02607.x] [Medline: 19320785]
24. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007;2007:26-30 [FREE Full text] [Medline: 18693791]
25. Cook DA, Teixeira MT, Heale BS, Cimino JJ, Fiol GD. Context-sensitive decision support (infobuttons) in electronic health records: a systematic review. J Am Med Inform Assoc 2017;24(2):460-468 [FREE Full text] [doi: 10.1093/jamia/ocw104] [Medline: 27497794]
26. Légat L, Van Laere S, Nyssen M, Steurbaut S, Dupont AG, Cornu P. Clinical decision support systems for drug allergy checking: systematic review. J Med Internet Res 2018;20(9):e258 [FREE Full text] [doi: 10.2196/jmir.8206] [Medline: 30194058]
27. Isbanner S, O'Shaughnessy P, Steel D, Wilcock S, Carter S. The adoption of artificial intelligence in health care and social services in Australia: findings from a methodologically innovative national survey of values and attitudes (the AVA-AI Study). J Med Internet Res 2022;24(8):e37611 [FREE Full text] [doi: 10.2196/37611] [Medline: 35994331]
28. Yeung YYK, Chen PQ, Ng PHF, Cheng ASK. Evaluation of the accuracy of the smart work injury management (SWIM) system to assist case managers in predicting the work disability of injured workers. J Occup Rehabil 2024. [doi: 10.1007/s10926-024-10199-7] [Medline: 38874680]
29. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
30. Hoel V, Feunou CM, Wolf-Ostermann K. Technology-driven solutions to prompt conversation, aid communication and support interaction for people with dementia and their caregivers: a systematic literature review. BMC Geriatr 2021 Mar 04;21(1):157. [doi: 10.1186/s12877-021-02105-0] [Medline: 33663401]
31. Rossen B, Lok B. A crowdsourcing method to develop virtual human conversational agents. Int J Hum-Comput Stud 2012;70(4):301-319. [doi: 10.1016/j.ijhcs.2011.11.004]
32. Shan Y, Ji M, Xie W, Qian X, Li R, Zhang X, et al. Language use in conversational agent-based health communication: systematic review. J Med Internet Res 2022;24(7):e37403 [FREE Full text] [doi: 10.2196/37403] [Medline: 35802407]
33. OpenAI. Introducing GPT-4o and more tools to ChatGPT free users. URL: https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/ [accessed 2024-05-20]
34. Cheng ST, Lau RWL, Mak EPM, Ng NSS, Lam LCW. Benefit-finding intervention for Alzheimer caregivers: conceptual framework, implementation issues, and preliminary efficacy. Gerontologist 2014;54(6):1049-1058. [doi: 10.1093/geront/gnu018] [Medline: 24688081]
35. Cheng ST, Lau RW, Mak EP, Ng NS, Lam LC, Fung HH, et al. A benefit-finding intervention for family caregivers of persons with Alzheimer disease: study protocol of a randomized controlled trial. Trials 2012;13(1):98 [FREE Full text] [doi: 10.1186/1745-6215-13-98] [Medline: 22747914]
36. Cheng ST, Fung HH, Chan WC, Lam LCW. Short-term effects of a gain-focused reappraisal intervention for dementia caregivers: a double-blind cluster-randomized controlled trial. Am J Geriatr Psychiatry 2016;24(9):740-750. [doi: 10.1016/j.jagp.2016.04.012] [Medline: 27401052]
37. Cheng ST, Chan WC, Lam LC. Long-term outcomes of the benefit-finding group intervention for Alzheimer family caregivers: a cluster-randomized double-blind controlled trial. Am J Geriatr Psychiatry 2019;27(9):984-994. [doi: 10.1016/j.jagp.2019.03.013] [Medline: 31076215]
38. Cheng ST, Mak EPM, Kwok T, Fung H, Lam LCW. Benefit-finding intervention delivered individually to Alzheimer family caregivers: longer-term outcomes of a randomized double-blind controlled trial. J Gerontol B Psychol Sci Soc Sci 2020;75(9):1884-1893. [doi: 10.1093/geronb/gbz118] [Medline: 31556447]

39. Cheng ST, Mak EPM, Fung HH, Kwok T, Lee DTF, Lam LCW. Benefit-finding and effect on caregiver depression: a double-blind randomized controlled trial. J Consult Clin Psychol 2017;85(5):521-529. [doi: 10.1037/ccp0000176] [Medline: 28287803]
40. Pot AM, Gallagher-Thompson D, Xiao LD, Willemse BM, Rosier I, Mehta KM, et al. iSupport: a WHO global online intervention for informal caregivers of people with dementia. World Psychiatry 2019;18(3):365-366 [FREE Full text] [doi: 10.1002/wps.20684] [Medline: 31496096]
41. Murray Alzheimer ResearchEducation Program. By Us For Us Guides. URL: https://the-ria.ca/resources/by-us-for-us-guides/ [accessed 2024-02-06]
42. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
43. Wu K, Su Y, Chu F, Chen AT, Zaslavsky O. Behavioral change factors and retention in web-based interventions for informal caregivers of people living with dementia: scoping review. J Med Internet Res 2022;24(7):e38595 [FREE Full text] [doi: 10.2196/38595] [Medline: 35797100]
44. Cheng ST, Lam LCW, Kwok T, Ng NSS, Fung AWT. The social networks of Hong Kong Chinese family caregivers of Alzheimer's disease: correlates with positive gains and burden. Gerontologist 2013;53(6):998-1008. [doi: 10.1093/geront/gns195] [Medline: 23371974]
45. Lim J, Goh J, Chionh HL, Yap P. Why do patients and their families not use services for dementia? Perspectives from a developed Asian country. Int Psychogeriatr 2012;24(10):1571-1580. [doi: 10.1017/s1041610212000919]

Abbreviations

AI: artificial intelligence
PDC30: Positive Dementia Caregiving in 30 Days

Edited by Y Jiang; submitted 27.06.24; peer-reviewed by P Ramaswamy, Y Shan; comments to author 22.09.24; revised version received 01.11.24; accepted 12.11.24; published 06.01.25.

Please cite as:
Cheng ST, Ng PHF
The PDC30 Chatbot—Development of a Psychoeducational Resource on Dementia Caregiving Among Family Caregivers: Mixed Methods Acceptability Study
JMIR Aging 2025;8:e63715
URL: https://aging.jmir.org/2025/1/e63715
doi:10.2196/63715
PMID:

Original Paper

Designing a Multimodal and Culturally Relevant Alzheimer Disease and Related Dementia Generative Artificial Intelligence Tool for Black American Informal Caregivers: Cognitive Walk-Through Usability Study

Cristina Bosco[1], MSc; Ege Otenen[1], MSc; John Osorio Torres[1], MSc; Vivian Nguyen[1], BSc; Darshil Chheda[1], MSc; Xinran Peng[1], MSc; Nenette M Jessup[2], MPH; Anna K Himes[2], BS; Bianca Cureton[2], BSPH; Yvonne Lu[2], PhD; Carl V Hill[3], PhD; Hugh C Hendrie[4], MD; Priscilla A Barnes[5], PhD; Patrick C Shih[1], PhD

[1]Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, United States

[2]School of Nursing, Indiana University, Indianapolis, IN, United States

[3]Alzheimer's Association, Chicago, IL, United States

[4]School of Medicine, Indiana University, Indianapolis, IN, United States

[5]School of Public Health, Indiana University, Bloomington, IN, United States

Corresponding Author:
Cristina Bosco, MSc
Luddy School of Informatics, Computing, and Engineering
Indiana University
700 N Woodlawn Ave
Bloomington, IN, 4740
United States
Phone: 1 812 856 5754
Email: cribosco@iu.edu

Abstract

Background: Many members of Black American communities, faced with the high prevalence of Alzheimer disease and related dementias (ADRD) within their demographic, find themselves taking on the role of informal caregivers. Despite being the primary individuals responsible for the care of individuals with ADRD, these caregivers often lack sufficient knowledge about ADRD-related health literacy and feel ill-prepared for their caregiving responsibilities. Generative AI has become a new promising technological innovation in the health care domain, particularly for improving health literacy; however, some generative AI developments might lead to increased bias and potential harm toward Black American communities. Therefore, rigorous development of generative AI tools to support the Black American community is needed.

Objective: The goal of this study is to test Lola, a multimodal mobile app, which, by relying on generative AI, facilitates access to ADRD-related health information by enabling speech and text as inputs and providing auditory, textual, and visual outputs.

Methods: To test our mobile app, we used the cognitive walk-through methodology, and we recruited 15 informal ADRD caregivers who were older than 50 years and part of the Black American community living within the region. We asked them to perform 3 tasks on the mobile app (ie, searching for an article on brain health, searching for local events, and finally, searching for opportunities to participate in scientific research in their area), then we recorded their opinions and impressions. The main aspects to be evaluated were the mobile app's usability, accessibility, cultural relevance, and adoption.

Results: Our findings highlight the users' need for a system that enables interaction with different modalities, the need for a system that can provide personalized and culturally and contextually relevant information, and the role of community and physical spaces in increasing the use of Lola.

Conclusions: Our study shows that, when designing for Black American older adults, a multimodal interaction with the generative AI system can allow individuals to choose their own interaction way and style based upon their interaction preferences and external constraints. This flexibility of interaction modes can guarantee an inclusive and engaging generative AI experience.

(JMIR Aging 2025;8:e60566) doi:10.2196/60566

KEYWORDS

multimodality; artificial intelligence; AI; generative AI; usability; black; African American; cultural; Alzheimer's; dementia; caregivers; mobile app; interaction; cognition; user opinion; geriatrics; smartphone; mHealth; digital health; aging

Introduction

Background

Alzheimer disease and related dementias (ADRD) are brain health conditions impacting cognitive abilities such as memory, language, problem-solving, and executive functions [1,2]. Within the United States, older Black American and Black adults have been identified as being twice as likely as individuals from other ethnic groups to develop ADRD over the course of their lifetime [3-5]. In a study conducted with 59,555 individuals, Black American adults had a 65% greater risk of developing ADRD (with 26.6/1000 Black American adults developing ADRD, compared with rates of 19.6/1000 for Latinos, 19.6/1000 for Pacific Islanders, 19.3/1000 for Whites, and 15.2/1000 for Asian Americans) [6].

However, despite being significantly more likely to develop ADRD, Black Americans exhibit lower levels of ADRD-related health literacy compared with other demographic groups [7]. This disparity impacts informal caregivers who often find themselves unprepared and uninformed about ADRD treatment and diagnosis, as well as about ADRD-related sources and help [8]. Therefore, providing essential medical information to these caregivers is crucial in combating the disease.

Considering the potential of new technologies to aid patients and caregivers with understanding medical language by simplifying content and removing jargon [9], generative artificial intelligence (AI) tools, like ChatGPT developed by OpenAI, could offer innovative opportunities in health care, particularly for improving ADRD-related health literacy. Indeed, recent studies proposed using ChatGPT for reliable and comprehensive health information [10-13].

However, relying solely on generative AI to improve underrepresented communities' health literacy can lead to several issues like providing inaccurate or false health information [14,15]. Additionally, studies have shown that using generative AI exclusively for health diagnosis and treatment can reinforce sexist and racial stereotypes and biases [16-18].

For instance, predictive models trained on historical data that lack representation from diverse demographics can reinforce existing biases in diagnosis and treatment plans [19,20]. The exclusion of certain populations in the studies and design process can lead to misdiagnosis and inappropriate treatment [21-23], exacerbating health inequalities rather than reducing them [20,24,25]. Therefore, it is crucial to ensure greater diversity in model training data and include diverse participants in studies and design processes to improve the generalizability of findings [26] and the fit of the technology use for the target population.

Therefore, given the potential of generative AI to improve health literacy and mitigate health inequalities among underrepresented populations, such as the Black American community, it is imperative to design inclusive ADRD health literacy interventions.

Previous studies, especially in the domain of human-centered design, highlight the importance of integrating cultural insights into health care technologies for this community [27-31]. Hence, it is crucial to tailor generative AI–based health interventions to meet the unique needs of Black American informal caregivers of individuals with ADRD.

Thus, this study identified the needs of Black American caregivers of individuals with ADRD by testing Lola, an ADRD-focused generative AI tool. Lola is a multimodal mobile app that provides access to health information from the Alzheimer's Association online repository [1] via text, chatbot, and voice assistant.

Related Work

Generative AI in Health Care

Generative AI has become a transformative force in health care, offering innovative solutions in diagnostics, personalized medicine, and patient care. Recent studies have highlighted the role of generative models in enhancing diagnostic accuracy by analyzing complex medical data and generating synthetic patient data to train other AI systems. For example, a study by Chen and Esmaeilzadeh [32] demonstrated that generative models like generative adversarial networks could significantly improve the detection and diagnosis of diseases by creating high-quality medical images from existing data. Additionally, variational autoencoders have been used to generate realistic medical imaging data, further aiding disease research and model training. Moreover, the use of large language models (LLMs) such as GPT-4 has shown potential for assisting clinicians by generating comprehensive patient histories and suggesting personalized treatment plans [33].

Moreover, the use of LLMs such as generative pretrained transformer (GPT)-4 has shown potential for assisting clinicians by generating comprehensive patient histories and suggesting personalized treatment plans [33]. Studies have also explored the use of LLMs in generating patient education materials and automating routine documentation tasks [34]. Despite still being in its developmental stages, generative AI holds significant promise for enhancing the accessibility of health care [10-13].

Human-Centered Generative AI

Generative AI models are trained on large data sets that may contain biases, further perpetuating stereotypes and disparities in health care [20,24]. More diversity is needed in both data used to train AI models and participants included in studies to ensure the generalizability of findings [26]. Additionally, addressing ethical and regulatory considerations is crucial when deploying AI in health care.

The integration of human-centered design principles into generative AI applications is crucial to ensure these technologies are accessible, intuitive, and beneficial to end users. Research has emphasized the importance of involving end users in the development process to create AI systems that meet their needs

and preferences. For instance, a study explored the application of ChatGPT for developing digital health interventions, finding that user involvement in the design phase led to higher satisfaction and engagement rates among patients [35]. Specifically, user feedback was gathered through focus groups and iterative testing phases, allowing developers to continuously refine the AI user interface and functionality.

Generative AI and the African American and Black Community

Inequalities have been observed in health care and accessibility to these services for underrepresented communities [36-40]. AI systems that initially aim to reduce the disparities in health care also present challenges for health care accessibility for Black American communities.

Addressing the unique health care challenges faced by Black American communities through generative AI is a growing area of interest. It is essential to ensure that these AI systems do not perpetuate existing biases or disparities. Studies have highlighted the potential of generative AI to provide culturally sensitive health care solutions. For example, a study explored the use of mobile health [41] interventions to mitigate the impact of COVID-19 in Black American communities, emphasizing the need for community-specific content and trusted communication channels to improve health outcomes.

Informal Caregivers of Individuals With ADRD and Their Use of Technology

In 2023, approximately 15 million people served as caregivers for individuals with ADRD [1], with African Americans accounting for one-tenth of these caregivers [1]. This unpaid work amounted to an estimated 16 billion hours [1]. Caregivers often feel unprepared for their roles [42], lacking information on diagnosis, treatment, and available resources [8,43].

Technology has been crucial in supporting these caregivers [44], and technologies have been developed to reduce caregiver stress and anxiety [45] and facilitate access to information for caregivers [46]. However, these technologies targeted different demographics than African American and Black populations. For instance, Hong et al [45] developed and tested a technological intervention aimed at reducing the stress and anxiety related with caregiving by focusing exclusively on Chinese participants. On a similar note, Boutilier et al [46] developed and tested a web app to help informal caregivers document information on their ADRD-diagnosed patients and share important changes with the care team. Boutilier et al [46] tested this system with mainly non-Hispanic White participants, with only 1% of participants being Black Americans.

However, no study has focused on Black American caregivers, which is extremely important and challenging at the same time, considering this population's lower level of technology use and adoption [47]. Additionally, a recent systematic review of technological interventions (mainly mobile apps) [48] found that most of these systems did not present any culturally relevant health information and were characterized by low accessibility and low readability.

Previous studies have shown that online health communities (OHCs) offer emotional and instrumental support to caregivers of individuals with ADRD [49] and chronic diseases [50], providing a platform to share experiences and find support from peers [49,51]. Furthermore, only a few technological support interventions for informal caregivers focus on ADRD-related information. Most aim to enhance caregivers' psychological well-being by providing stress management strategies [52] or developing social games to promote physical activity and mutual support networks [53].

We identified 2 gaps in the previous literature. Although OHCs offer instrumental support by sharing caregiving tips, they often lack localized, actionable ADRD information or personalized support tailored to caregivers' and patients' specific needs. Additionally, there is limited representation of Black American caregivers in research, raising concerns about the ability to understand and address their sociocultural background as well as their unique needs compared with other populations, which is vital for developing health technologies for them [49].

Methods

Recruitment

We conducted recruitment and research sessions in close collaboration with a community engagement team who ensured that participants were all caregivers in some capacity for individuals living with ADRD and fulfilled the inclusion criteria. To meet the inclusion criteria, participants had to be older than 50 years, identify as Black Americans, reside in the target area, and actively provide care for a person with ADRD. Throughout this study, emphasis was placed on collaboration with the Black American community within the region. Consequently, 2 community engagement specialists joined the research team to organize monthly community advisory team board meetings to maintain community involvement and awareness of research insights and findings. Prior to implementing any major design decisions, the advisory team board was consulted to ensure cultural relevance. These measures enabled the research team to make more culturally competent decisions.

Ethical Considerations

Ethical approval was provided by the institutional review board at Indiana University (approval number 12241), All participants received a full explanation of the consent procedures, and any specials needs were accommodated during sessions. Participants were compensated with a US $50 gift card for taking part. Participants' privacy and confidentiality were protected by anonymizing the data, changing the wording of the quotes presented in the Results section, and deleting all the audio recordings after the transcriptions were completed.

Data Collection and Analysis

We performed an adapted cognitive walkthrough with users [54,55]. We conducted 15 sessions in 2 regions in the Midwest region of the United States. Each session was completed in a 1-on-1 meeting that involved 1 researcher and 1 notetaker with the participant. Each session lasted approximately 30 minutes. Although we were aware that longer testing sessions might guarantee the collection of more comprehensive feedback, we

made sure to construct the tasks and testing in a way that allowed the emergence of in-depth insights and findings. Moreover, given the older age of our participants, we did not want to induce cognitive fatigue.

During a session, each participant was introduced to the platform and a set of 3 tasks. Those 3 tasks consisted of searching for an article about brain health, searching for ADRD-related local events, and finally, searching for opportunities to participate in scientific research.

As participants performed the tasks, they were asked to voice their thoughts out loud in real time ("think aloud" method [56]), and after each task was complete, there was a brief semistructured discussion centered around 4 tenets: (1) usability (eg, Did you find anything in this task confusing? In what way?), (2) accessibility (eg, How do you expect Lola to answer you? Did you find the information easy to understand? If not, why?), (3) cultural relevance (eg, Did you find the tone of interactions appropriate? If not, how so? How do you think this platform can be made better for the Black American community?), (4) adoption (eg, What do you think would be a good way of spreading knowledge about this so that people you know can properly adopt it and use it?)

Each session was audio-recorded and transcribed. The generated transcripts were spell-checked. Guided by the principles of usability, accessibility, and relevance, 6 researchers coded the transcripts using the open-source software Taguette [57]. The team took an analytical, iterative approach by discussing the codes and consolidating them into themes [58] through second and third rounds of discussion. The resulting major themes about the needs of the users are described in the Results section.

System Design

Overview

Generative AI is at the forefront of technological advancements in health care applications, specifically in the development of interactive tools such as chatbots and voice assistants. In the context of Lola, an innovative app tailored for Black Americans coping with ADRD, generative AI plays a crucial role. It powers both the chatbot and voice assistant, enabling them to deliver intuitive and empathetic user interactions. These AI-driven tools analyze complex user queries, source relevant information, and articulate responses that are both informative and comforting.

Integration With ChatGPT Technology

The core functionalities of Lola—the chatbot and voice assistant—use the natural language processing prowess of ChatGPT 3.5 Turbo.

Voice Assistant and Chatbot Mechanism

These multimodal features enhance the accessibility of the app, providing users with options for both auditory and textual communication. This is particularly beneficial for users with different levels of technological literacy or physical abilities. The real-time response generation capability of ChatGPT ensures that every interaction is smooth and engaging.

Data Utilization and Information Sourcing

Lola's AI systems are designed to draw information from authoritative sources such as the Alzheimer's Association online repository [1], supplemented by curated recommendations from a team of medical doctors specializing in ADRD. This method ensures that the information provided is not only accurate but also up to date with current medical standards and practices.

Ensuring Clear and Meaningful Communication

The generative process involves a sophisticated workflow in which initial responses, structured using LLM techniques, are refined by ChatGPT for clarity and empathy. This ensures that the communication is not only technically accurate but also tailored to meet the emotional and cognitive needs of the users.

Technical Infrastructure

The comprehensive technical infrastructure of Lola uses Flutter for development of the front end to maintain a consistent and engaging user experience with visually appealing, high-performance interactions across various devices and modalities. The back end is powered by Python's Flask framework, paired with a PostgreSQL database for robust data management and fast data retrieval during the human-app interactions. The deployment of Docker for containerization, along with the use of the research team's university servers, provides a secure and reliable environment for hosting and managing Lola's functionalities. By reducing compatibility issues, this setup ensures consistent system performance and reliability.

App Features

In response to the challenges faced by older Black Americans regarding the accessibility of credible ADRD information, Lola's interface serves as a platform for accessing ADRD-related articles, local and virtual events, and opportunities to engage in the latest ADRD research. Lola incorporates a chatbot feature represented by a floating green bubble in the bottom right corner, accessible to users throughout their interaction with the app (Figure 1).

Users can interact with the Lola chatbot by clicking the green call-to-action button, which initiates a conversation through text or voice input. Using generative AI, the Lola chatbot provides responses in text format or as voice messages by clicking the "Play" button, facilitating accessibility for users with varying literacy levels or preferences for voice interaction. Users can also copy the response as text or share the response with their contacts. When users choose to end the conversation, the system saves the conversation in the chat history (Figure 2).

Figure 1. Example of a user having a conversation with the Lola chatbot using the interface designed by the research team.

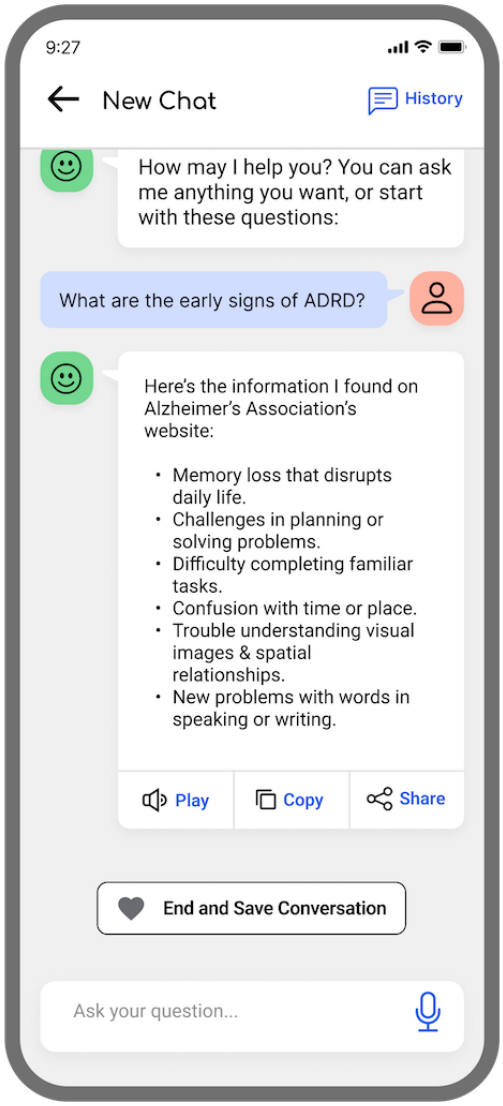
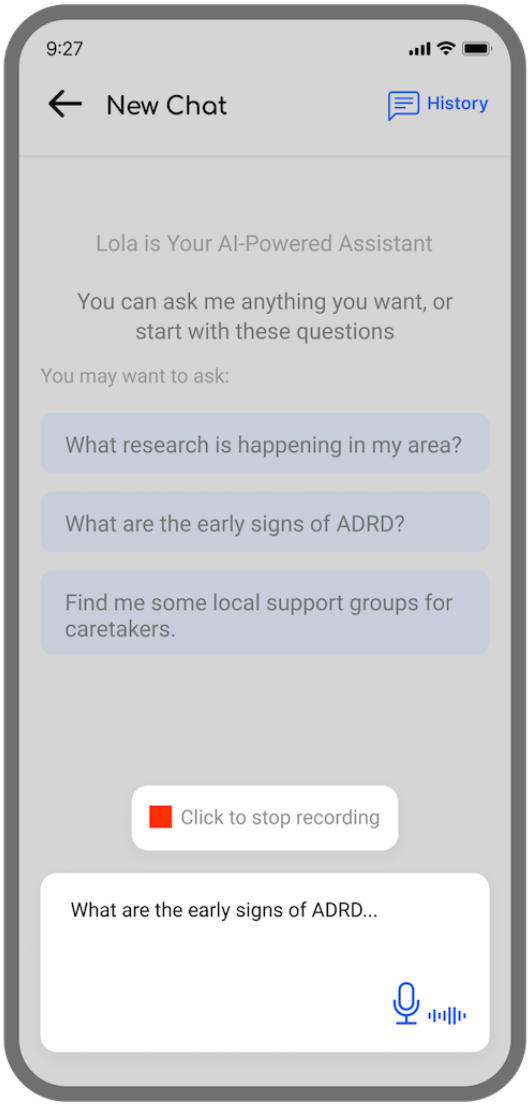


Figure 2. Example of a user using voice input and the possibility of browsing the history.



Results

Participant Demographic Information

A total of 15 participants were included in this study. All of them were informal caregivers for someone living with ADRD and were Black American. All participants shared that they were born and raised in the United States. Regarding family origins, 5 participants identified having African family origins, 9 were unsure of their family origins, and 1 chose not to disclose. Table 1 summarizes the participants' demographic data. Of the 15 participants, 12 (80%) were 65 years or older. Most participants were female (11/15, 73%), with 2 (2/15, 13%) being male and 2 (2/15, 13%) providing no gender response. Regarding education, 10 participants (10/15, 67%) indicated having at least some college education (partial completion, a bachelor's degree, or an associate degree), while 2 (2/15, 13%) had a graduate degree, 2 (2/15, 13%) indicated a high school education, and 1 (1/15, 7%) did not provide a response. Most participants (11/15, 73%) were retired, only 3 participants (3/15, 20%) worked either part-time or full-time, and 1 (1/15, 7%) did not respond.

Table 1. Participant demographic information.

Participant number	Age (years)	Gender	Education level	Employment	Work sector
1	72	Female	Some college	Retired	Health care
2	74	Male	Some college	Retired	Utilities
3	68	Female	Associate's degree	Retired	NR[a]
4	69	Male	Some college	Retired	Construction
5	54	Female	Associate's degree	Part-time	Health care
6	66	NR	NR	NR	NR
7	69	Female	Graduate degree	Retired	Community advocacy
8	69	Female	Some college	Retired	Health care
9	75	Female	Some college	Retired	Faith-based
10	71	Female	High school	Retired	Customer service
11	69	Female	Associate's degree	Retired	Insurance
12	72	NR	High school	Retired	Health care
13	68	Female	Some college	Retired	Health care
14	63	Female	Graduate degree	Full-time	Employment specialist
15	62	Female	Bachelor's degree	Part-time	Education

[a]NR: not reported.

Overview of Lola

The Lola system includes 3 different models of interaction: (1) accessing information in a traditional touch screen app format, (2) texting with an AI-powered text-based chatbot, and (3) talking to a voice assistant that accompanies the chatbot. Participants experienced all 3 models of interaction during user testing.

Themes

From the analysis of the data collected, 3 main themes emerged: (1) the need for a multimodal AI tool that relies on multiple senses, (2) the need for culturally specific ADRD content, and (3) the importance of physical places and existing social networks to spread Lola among the community. This section will explore these themes in depth and how they relate to the goals of the research.

Theme 1: The Need for a Multimodal AI Tool That Relies on Different Senses

This theme describes participants' need for interacting with the Lola system in different ways aside from the traditional touch screen gesturing predominant in most mobile apps. This multimodal interaction could take several forms, for example, voice interaction that can read content generated as text out loud. This would allow participants with low vision, which is common among older adults, to access the content.

> *Because with the people that I work with on a daily basis, one is not able to read anymore because she's going blind. So, it's not just Alzheimer's dementia that needs to be looked at, at least all the senses need to be looked at because people are going through so many different things at the same time.* [P4]

Consistently, auditory access to information positively influenced participants' interactions with the system. Some participants preferred to engage with the tool using their voice, hence treating the tool as a voice assistant and asking it questions. For the participants, this was a valuable addition to the traditional touch screen interaction. For people not only with low vision but also with lower digital skills, voice interaction resembles a conversational exchange with which they are familiar and comfortable.

> *I think that this was good to be able to speak it. Because, if you're dealing with people who are not computer savvy or internet savvy, being able to speak it, and it automatically pops up or speaks your question and then being able to close that area so it won't keep recording. That was quick and easy. It was simple.* [P9]

Speaking features (as opposed to just reading and typing information) are especially helpful since they would enable people to access their usual methods of dealing with health-related information, such as writing it down on a piece of paper and saving it for later reference while someone is sharing the information. Having an auditory feature to provide health care information in addition to other interaction models can improve the accessibility of relevant information for those with difficulties accessing digital content due to low visual capacity.

> *That would help too if it played. And then they spoke out the information. And there's different ways I can have a pen and a pad. And then while she's speaking, I could write the information down. Because it is printed small. And you have to read all this information. But if she tells me, and the number you can call is one 800, I'd be writing it down.* [P1]

However, solely relying on the sense of hearing could be a limitation for people with auditory impairments, who could get frustrated by the tool.

> *But I couldn't hear it. I was kind of frustrated. So, it kind of forced me to read it in my eyes are strained.* [P14]

One participant highlighted the benefit of a multimodal approach pairing visual and auditory stimuli together, drawing on their experience in church where reading and listening occur simultaneously.

> *For me and people lazy, they don't want to have to go through the reading it would have, because I would like to hear it while I'm reading it. It is better. If it were just speaking it, I probably wouldn't grasp. But it was speaking in I was reading it and I was falling through. It would have been like being in church. And the preacher is preaching a sermon, he tell you what he coming out of and you fall you open your Bible to that. John 316, and you follow him as he's talking.* [P5]

This theme highlights the importance of creating a system with multimodal interaction and pairing auditory and visual inputs in a way that makes the tool more accessible and allows a broader audience to meaningfully engage.

Theme 2: The Need for Context and Culture-Specific ADRD Content

This theme highlighted the need for contextually and culturally specific mobile apps that integrate relevant and accessible ADRD information. When combined with a multimodal AI tool, these tailored apps can significantly enhance participants' access to the information they need, which is not always directly related to health care.

One participant expressed their discontent with the lack of relevant information in the Lola app, stating that they wanted more than just medical information such as actionable lifestyle changes that can improve their health.

> *Definitely more than just medical information. Because I'm not one to push drugs, I don't take none. If a doctor tells me something is going on, I'm asking first and foremost, what can I do? Or how can you help me get to this process? Naturally? Do I need to implement exercise, do I need to change my diet, what vitamins, whatever. But if they talk about medication, it's gonna be a big turn off for me instantly. So that's just me. Some people think a quick fix is the answer. But every Quick Fix has consequences.* [P4]

Additionally, another participant mentioned the need for accessible ADRD mobile apps. Being able to operate and navigate through the app with little or without outside help is ideal.

> *Well, you know I just had therapy at the hospital, okay. And I was so amazed that every time I went to therapy, they gave me a folder with exercises, and one time they gave me an app. A girl was there, and she showed that to me. And I was just shocked. She used the app to show me the exercises, and every time I went back, she added exercises for me to do on the app. So that was amazing to me. So yeah, the apps are important. They really are. And they're to be simple. And they got to be informative. And they got to be to the point where even the agent can navigate through them.* [P10]

Some participants expressed discontent with the use of language in the Lola app, as it was perceived as too technical and not catering to the target population well enough.

> *Now all these big words, that's what we get me. I wouldn't understand them. Okay, so there's too many words, too many big words, I would not understand.* [P15]

Certain participants expressed a desire for the mobile app to cater to their Black American communities by offering culturally specific ADRD information tailored to their needs. Consequently, they sought additional content that would address their unique concerns and circumstances.

> *So I'm thinking on the other side as the creators of it, maybe you don't want to offend other races or whatever. But if Alzheimer's is more prevalent in my community, I feel like we need to have that target. So we will know exactly where to go, filter through some of the other stuff, some of the things but Alzheimer's and dementia are across the board. But maybe some things are just specific toward my race. Sure. And I want to know that, and I want to know that in the most timely fashion possible.* [P9]

Last, some of the participants highlighted the necessity for the mobile app to include community-centric information and information about local resources and organizations for ADRD, which could be tailored to individual users and their community based on their location.

> *I don't travel too far for it. So, I think there needs to be something located in their community.* [P4]
>
> *More websites because a lot of people are interested in a study of Alzheimer's or dementia. And it's national, local. So, I would rather you know, her give me the local if I had to visit it. It's good to know the national, but also there's local organizations as well.* [P1]

This section showed how providing context and culturally relevant ADRD information that is accessible is requested by the Black American community.

Theme 3: Physical Places and Existing Social Networks Are Necessary to Spread Lola Among the Community

This last theme is related to the importance of using physical community-oriented places and pre-existing social networks to spread technological tools among the Black American community. Our participants suggested community-oriented ways to advertise these health-related technologies to be adopted and used by the community. Most of the participants mentioned the importance of doing demonstrations of the Lola app in physical community places such as a town hall.

> *I think if they're going to implement this process, then they should have maybe a town hall meeting to go through the process have a big screen. The people that are interested in finding out more than they'll do it step by step. So I think that will be a demonstration we're having at a community center or, or even in a building like this, and advertise that we have a new application now called Lola that will help.* [P1]

Finally, participants mentioned how, in the advertisement of the Lola app, it is important to leverage pre-existing social networks within the community, such as relying on word of mouth and showing how people within the community are using Lola.

> *More information to let us know, what makes Lola different from the other? That will make me want to go to Lola more. And that could cover word of mouth, that could cover the internet, having more of these to get people involved to know more, maybe part of the marketing or the draw from my community would be Lola specifically, or is more strategic as far as getting information out to Black communities.* [P9]

This section highlighted how physical locations where the community gathers and existing social networks are vital to promote increased use of Lola, as word-of-mouth marketing is effective when performed by trustworthy individuals.

Discussion

Principal Findings

Deployment of generative AI in the health care domain offers the promise of innovation and expands the potential of equitable health technologies. However, erroneous design and development of generative AI could increase bias and potential harm toward historically marginalized and underserved communities. Including underserved populations in the design and development of a technological intervention has been shown to increase the opportunity for representation and inclusivity [27-31]. Following this line of research, our study investigated the experiences of Black American caregivers using a generative AI tool to seek ADRD-related information (such as symptoms, community resources, financial aid).

This is the first study that specifically addresses the needs, demands, and challenges of this specific population from the perspective of caregivers of individuals with ADRD. An in-depth description of the 3 key design implications to create an inclusive and culturally relevant AI-supported platform for health will follow.

Fostering Inclusivity Through Multimodality

We observed how Black American caregivers expressed the need for multimodality of interaction to equitably access what generative AI tools have to offer. Indeed, they expressed the importance of being able to "talk with the tool" and hear back (eg, read content out loud) in the fashion of a conversational voice assistant. Several participants' insights suggested a preference toward multimedia, emphasizing the auditory and visual senses and providing modes (speaking, reading, and listening) to support these interaction styles. These implications may originate from their day-to-day time constraints, requiring short and efficient interactions.

Previous work addressed the efficiency of using multimodal systems that provide health care. Linders et al [59] created a multimodal health assistant to provide health information. Soubutts et al [60] used a multimodal device to support the needs of households for health and care purposes during COVID-19. The need for multimodal interactions for memory aid systems has been also highlighted [61].

Additionally, multimodality can guarantee a larger audience who will interact with the system [62], thereby allowing people with low vision and low hearing to access the system and seek ADRD health-related information. Within the landscape of designing for inclusive and equitable computing, our study took the approach of multimodal interaction. Our findings emphasize the need to tailor technology to accommodate different interaction styles. This can be achieved by designing multimodal technologies that can meet the needs of differently abled users.

Answering the Call for Context and Culture-Specific ADRD Information

Our findings reflected the caregivers' need for contextually and culturally relevant ADRD information that is specific to their communities or personalized based on their needs and preferences. They indicated their need for lifestyle tips and actionable tools for ADRD rather than only medical information in the system. They also did not want to be offered only general ADRD medical information; rather, they wanted ADRD-related information specific to their community, local resources, and geography. Locality and cultural relevance emerged as 2 of the most important factors influencing the perceived usefulness of the information sought. This is in line with previous studies [63,64] showing that providing community-centered information can increase the acceptance and adoption of generative AI tools. Compared with the use of OHC, which provides general emotional and social support, as identified by previous studies [49-51], our study reveals the need for specialized information on health care. Our participants expressed their need for instrumental and informational support specific to their social and cultural context as well as geographic location. This finding complements the benefits of OHCs to meet all the needs of caregivers including instrumental and information support.

Furthermore, our participants wanted to be sure that the information provided was accessible to people with low literacy skills within their own communities. This is in line with the collectivist nature of African American communities observed in a prior study on technology use [30], wherein, for a technological intervention to be accepted, it must be inclusive of all the members of the Black American community.

Keeping Not Only the People But Also the Community in the Loop

Finally, our study offers a novel contribution in generative AI research investigating how generative AI tools can be advertised and spread among the Black American community. To our knowledge, this is the only study that has investigated this aspect of generative AI use and adoption. Indeed, low technology adoption is a striking problem within Black American

communities, especially among people with low economic status, which prevents them from using online tools to access health information and improve their health literacy [65]. However, low adoption does not arise from the lack of digital tools but also from a skepticism toward digital tools [66]. Privacy is also a crucial concern, as Black American populations tend to be more skeptical toward the use of technology based on a fear of privacy intrusions and breaches of privacy [41,63]; thus, generative AI–based technologies for this population should foster a transparent and reliable way of handling user privacy.

Some studies [35,66-68] have investigated how faith-based organizations can be important for health dissemination and assisting Black American adults seeking health information. Our study expands on these previous studies by exploring the strategies suggested by our participants, which could be used to foster generative AI adoption and its use among their communities.

Hence, although most studies have focused on programs designed to disseminate information based upon the creation of seminars and physical events, our study provides evidence of how physical organizations (faith-based organizations, public as well as medical organizations) could be strategic actors in fostering the adoption and use of a generative AI tool for ADRD. Additionally, given that these populations have limited resources and opportunities to interact with AI systems and often have difficulties accessing any technology (eg, software, hardware, internet) due to financial and infrastructural challenges [39], incorporating these physical organizations is significant. Building trust and familiarity by involving faith-based organizations and communities in the dissemination of the tool, as well as in the tool's use, facilitates not only dissemination but also the ability to break the barrier of skepticism toward health technology, by showing the community how it could embrace the tool.

Finally, ethical implications and concerns must be addressed in the discussion of the findings, even though ethical concerns did not explicitly emerge in this study as its focus was on usability, accessibility, cultural relevance, and adoption. However, it is crucial to discuss how topics such as consent, privacy, and bias can impact the use of generative AI for an underrepresented community. As suggested by previous studies, ethical guidelines committed to social good and human well-being should be developed and followed by AI system developers [32]. Consent and data management must be at the core of every generative AI system, allowing users to be in full control of their own data and their interaction with the tool. Furthermore, it is critical that users who interact with or provide data for AI systems fully understand what they are consenting to. This is especially important for populations with lower literacy rates, for whom the clarity and accessibility of consent forms must be carefully considered.

Limitations

This study is subject to several limitations. First, the sample characteristics are not representative of the broader population. The age range of our participants may not adequately capture the unique challenges faced by younger caregivers who juggle multiple responsibilities alongside caregiving. Future studies should aim to recruit a more diverse sample to better understand the challenges different demographics face in accessing health information for caregiving.

Furthermore, our participants demonstrated a high level of education, with the majority (12/15, 80%) having some college education. This could have influenced our findings, potentially skewing the results toward those with higher educational backgrounds. We recommend future research include Black American individuals with lower literacy levels to better understand their specific needs and how they engage with AI-generated health technologies.

Last, the user experience testing sessions for this study were conducted outside of the participants' daily environments, which may not accurately reflect the day-to-day challenges they might encounter when interacting with Lola. This limitation could restrict the scope of identified problems. Future studies should consider conducting tests in settings and scenarios that more closely mimic real-life conditions to gain deeper insights into the tool's practical efficacy and user engagement.

Apart from the opportunities already identified, our study opens the research space on the intersection between generative AI and underrepresented populations living in the United States. One possibility could be investigating the effectiveness of tailoring ADRD content to the community's values; hence, future studies could quantify the effectiveness of such interventions on user engagement as well as on ADRD-related health literacy. Second, future studies could investigate the impact of different interaction modalities on user engagement, thus understanding whether multimodality is preferred over solely having the opportunity of interacting with one modality, and explore the potential benefit of using different modalities over others in groups with various needs.

Conclusions

This study adds to the expanding field of research focused on deploying generative AI to improve health literacy. More specifically, the emphasis of this study was on improving the ADRD-related health literacy of Black American informal caregivers, who often are solely responsible for individuals with ADRD. Our objective in testing our intervention, Lola, was to identify the specific aspects of our design that need to be tailored and modified. This approach aims to develop a tool that can genuinely assist this population, enhance their ADRD-related health literacy, and support them in their caregiving responsibilities.

Our findings showed that ADRD-specific content should be made contextually and culturally relevant to be perceived as useful by informal caregivers, thereby including information on local, geographically dependent resources. This shows how generative AI designed to improve ADRD-related health literacy should be highly personalized and dependent upon the user's social and cultural context. Additionally, concerning the interactions, relying on multimodality can contribute to creating an accessible tool for informal caregivers, by offering them multiple and diverse interaction strategies. Once again, this

finding indicates the need for a generative AI tool that can adapt to the users' physical and cognitive needs and demands.

Finally, through active collaboration and community engagement, this study provides evidence of the importance of including the users and population of interest in the evaluation of a design as a way to guarantee that generative AI technologies will benefit all individuals within society, even underrepresented communities such as Black American populations.

Acknowledgments

Funding was provided by the National Health Institute (NIA 1R24AG071471) and the National Science Foundation (NSF number 2145049).

Authors' Contributions

CB and EO wrote the majority of the manuscript and led the writing process and some of the detailed cognitive walk-through sessions. EO, JOT, VN, DC, and XP wrote and edited portions of the manuscript and led some of the detailed cognitive walk-through sessions. BC, AKH, and NMJ led participant recruitment and community engagement. PAB, YL, CVH, HCH, and PCS initiated and ideated the project.

Conflicts of Interest

None declared.

References

1. Alzheimer's Disease Facts and Figures. Alzheimer's Association. URL: https://www.alz.org/alzheimers-dementia/facts-figures [accessed 2024-12-21]
2. Shojaei F. Uncovering the Impact of Co-Design and Design Thinking in Enhancing the Empowerment of People with Dementia. 2024 Presented at: International Conference on Innovation of Emerging Communication and Information Technology; October 20-23, 2024; Casablanca, Morocco.
3. Barnes LL. Alzheimer disease in African American individuals: increased incidence or not enough data? Nat Rev Neurol 2022 Jan;18(1):56-62 [FREE Full text] [doi: 10.1038/s41582-021-00589-3] [Medline: 34873310]
4. Barnes LL, Bennett DA. Alzheimer's disease in African Americans: risk factors and challenges for the future. Health Aff (Millwood) 2014 Apr;33(4):580-586 [FREE Full text] [doi: 10.1377/hlthaff.2013.1353] [Medline: 24711318]
5. Brewster P, Barnes L, Haan M, Johnson JK, Manly JJ, Nápoles AM, et al. Progress and future challenges in aging and diversity research in the United States. Alzheimers Dement 2019 Jul 18;15(7):995-1003 [FREE Full text] [doi: 10.1016/j.jalz.2018.07.221] [Medline: 30240574]
6. Mayeda ER, Glymour M, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. Alzheimers Dement 2016 Mar 10;12(3):216-224 [FREE Full text] [doi: 10.1016/j.jalz.2015.12.007] [Medline: 26874595]
7. Muvuka B, Combs RM, Ayangeakaa SD, Ali NM, Wendel ML, Jackson T. Health literacy in African-American communities: barriers and strategies. Health Lit Res Pract 2020 Jul 16;4(3):e138-e143. [doi: 10.3928/24748307-20200617-01] [Medline: 32674161]
8. Xie B, Champion JD, Kwak J, Fleischmann KR. Mobile health, information preferences, and surrogate decision-making preferences of family caregivers of people with dementia in rural Hispanic communities: cross-sectional questionnaire study. J Med Internet Res 2018 Dec 10;20(12):e11682 [FREE Full text] [doi: 10.2196/11682] [Medline: 30530450]
9. Ayre J, Mac O, McCaffery K, McKay BR, Liu M, Shi Y, et al. New frontiers in health literacy: using ChatGPT to simplify health information for people in the community. J Gen Intern Med 2024 Mar;39(4):573-577 [FREE Full text] [doi: 10.1007/s11606-023-08469-w] [Medline: 37940756]
10. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 08;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
11. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023 Feb;2(2):e0000198 [FREE Full text] [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
12. Samaan JS, Yeo YH, Rajeev N, Hawley L, Abel S, Ng WH, et al. Assessing the accuracy of responses by the language model ChatGPT to questions regarding bariatric surgery. OBES SURG 2023 Apr 27;33(6):1790-1796. [doi: 10.1007/s11695-023-06603-5]
13. Walker HL, Ghani S, Kuemmerli C, Nebiker CA, Müller BP, Raptis DA, et al. Reliability of medical information provided by ChatGPT: assessment against clinical guidelines and patient information quality instrument. J Med Internet Res 2023 Jun 30;25:e47479 [FREE Full text] [doi: 10.2196/47479] [Medline: 37389908]

14. Bahrini A, Khamoshifar M, Abbasimehr H, Riggs RJ, Esmaeili M, Majdabadkohne RM. ChatGPT: Applications, Opportunities, and Threats. 2023 Presented at: Systems and Information Engineering Design Symposium (SIEDS); April 27-28, 2023; Charlottesville, VA. [doi: 10.1109/sieds58326.2023.10137850]
15. Biswas SS. Role of Chat GPT in public health. Ann Biomed Eng 2023 May 15;51(5):868-869. [doi: 10.1007/s10439-023-03172-7] [Medline: 36920578]
16. Baldassarre MT, Caivano D, Fernandez Nieto B, Gigante D, Ragone A. The Social Impact of Generative AI: An Analysis on ChatGPT. 2023 Presented at: GoodIT '23: ACM International Conference on Information Technology for Social Good; September 6-8, 2023; Lisbon, Portugal. [doi: 10.1145/3582515.3609555]
17. McDuff D, Song Y, Kapoor A, Ma S. Characterizing bias in classifiers using generative models. 2019 Presented at: NIPS'19: 33rd International Conference on Neural Information Processing Systems; December 8-14, 2019; Vancouver, British Columbia, Canada.
18. Spector-Bagdady K. Generative-AI-generated challenges for health data research. Am J Bioeth 2023 Oct 09;23(10):1-5. [doi: 10.1080/15265161.2023.2252311] [Medline: 37831940]
19. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Jun 21;2(4):230-243. [doi: 10.1136/svn-2017-000101]
20. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. Future Internet 2023 Aug 24;15(9):286. [doi: 10.3390/fi15090286]
21. Bautista YJP, Theran C, Aló R, Lima V. Health Disparities Through Generative AI Models: A Comparison Study Using a Domain Specific Large Language Model. In: Lecture Notes in Networks and Systems. Cham, Switzerland: Springer Nature; 2023:220-232.
22. Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. Proc Natl Acad Sci U S A 2016 Apr 19;113(16):4296-4301 [FREE Full text] [doi: 10.1073/pnas.1516047113] [Medline: 27044069]
23. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health 2019 Dec;9(2):010318. [doi: 10.7189/jogh.09.020318] [Medline: 31788229]
24. Petti U, Nyrup R, Skopek JM, Korhonen A. Ethical considerations in the early detection of Alzheimer's disease using speech and AI. 2023 Presented at: FAccT '23: the 2023 ACM Conference on Fairness, Accountability, and Transparency; June 12-15, 2023; Chicago, IL. [doi: 10.1145/3593013.3594063]
25. Tamkin A, Brundage M, Clark J, Ganguli D. Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models. The Stanford Natural Language Processing Group. URL: https://www-nlp.stanford.edu/pubs/tamkin2021understanding.pdf [accessed 2024-12-21]
26. Mirkin S, Albensi BC. Should artificial intelligence be used in conjunction with neuroimaging in the diagnosis of Alzheimer's disease? Front Aging Neurosci 2023 Apr 18;15:1094233 [FREE Full text] [doi: 10.3389/fnagi.2023.1094233] [Medline: 37187577]
27. Dillahunt TR. Fostering social capital in economically distressed communities. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014; Toronto, Ontario, Canada. [doi: 10.1145/2556288.2557123]
28. Dillahunt TR, Kameswaran V, Li L, Rosenblat T. Uncovering the Values and Constraints of Real-time Ridesharing for Low-resource Populations. 2017 Presented at: CHI '17: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025470]
29. Harrington CN, Piper AM. Informing Design through Sociocultural Values: Co-Creation with Low-Income African-American Older Adults. 2018 Presented at: PervasiveHealth '18: 12th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 21-24, 2018; New York, NY. [doi: 10.1145/3240925.3240966]
30. Parker AG, Grinte RE. Collectivistic health promotion tools: Accounting for the relationship between culture, food and nutrition. International Journal of Human-Computer Studies 2014 Feb;72(2):185-206. [doi: 10.1016/j.ijhcs.2013.08.008]
31. Richards OK, Marcu G, Brewer RN. Hugs, Bible Study, and Speakeasies: Designing for Older Adults' Multimodal Connectedness. 2021 Presented at: DIS '21: Designing Interactive Systems Conference 2021; June 28-July 2, 2021; Virtual meeting. [doi: 10.1145/3461778.3462075]
32. Chen Y, Esmaeilzadeh P. Generative AI in medical practice: in-depth exploration of privacy and security challenges. J Med Internet Res 2024 Mar 08;26:e53008 [FREE Full text] [doi: 10.2196/53008] [Medline: 38457208]
33. Mesko B. The ChatGPT (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals. J Med Internet Res 2023 Jun 22;25:e48392 [FREE Full text] [doi: 10.2196/48392] [Medline: 37347508]
34. Rodriguez DV, Lawrence K, Gonzalez J, Brandfield-Harvey B, Xu L, Tasneem S, et al. Leveraging generative AI tools to support the development of digital solutions in health care research: case study. JMIR Hum Factors 2024 Mar 06;11:e52885 [FREE Full text] [doi: 10.2196/52885] [Medline: 38446539]
35. Knott CL, Miech EJ, Slade J, Woodard N, Robinson-Shaneman B, Huq M. Evaluation of organizational capacity in the implementation of a church-based cancer education program. Glob Implement Res Appl 2022 Jan 07;2(1):22-33. [doi: 10.1007/s43477-021-00033-0]

36. Etori N, Temesgen E, Gini M. What we know so far: artificial intelligence in African healthcare. arXiv. Preprint posted online on May 10, 2023 .
37. Ganju A, Satyan S, Tanna V, Menezes SR. AI for improving children's health: a community case study. Front Artif Intell 2020;3:544972 [FREE Full text] [doi: 10.3389/frai.2020.544972] [Medline: 33733204]
38. Guo J, Li B. The application of medical artificial intelligence technology in rural areas of developing countries. Health Equity 2018 Aug;2(1):174-181 [FREE Full text] [doi: 10.1089/heq.2018.0037] [Medline: 30283865]
39. Hoodbhoy Z, Hasan B, Siddiqui K. Does artificial intelligence have any role in healthcare in low resource settings? J Med Artif Intell 2019 Jul;2:13-13. [doi: 10.21037/jmai.2019.06.01]
40. Senteio C, Murdock PJ. The Efficacy of Health Information Technology in supporting health equity for Black and Hispanic patients with chronic diseases: systematic review. J Med Internet Res 2022 Apr 04;24(4):e22124 [FREE Full text] [doi: 10.2196/22124] [Medline: 35377331]
41. Harris KM, Mayo Gamble T, Yoo MG, Spell LA, Minor TN, Jones H, et al. Leveraging mHealth to mitigate the impact of COVID-19 in Black American communities: qualitative analysis. JMIR Hum Factors 2023 Dec 22;10:e47294 [FREE Full text] [doi: 10.2196/47294] [Medline: 37874735]
42. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues in Clinical Neuroscience 2022 Apr 01;11(2):217-228. [doi: 10.31887/dcns.2009.11.2/hbrodaty]
43. Wozney L, Freitas de Souza LM, Kervin E, Queluz F, McGrath PJ, Keefe J. Commercially available mobile apps for caregivers of people with Alzheimer disease or other related dementias: systematic search. JMIR Aging 2018 Dec 07;1(2):e12274 [FREE Full text] [doi: 10.2196/12274] [Medline: 31518255]
44. Bangerter LR, Griffin J, Harden K, Rutten LJ. Health information-seeking behaviors of family caregivers: analysis of the health information national trends survey. JMIR Aging 2019 Jan 14;2(1):e11237 [FREE Full text] [doi: 10.2196/11237] [Medline: 31518309]
45. Hong Y, Shen K, Han H, Ta Park V, Bagchi P, Lu H, et al. A WeChat-based intervention, Wellness Enhancement for Caregivers (WECARE), for Chinese American dementia caregivers: pilot assessment of feasibility, acceptability, and preliminary efficacy. JMIR Aging 2023 Apr 5;6:e42972. [doi: 10.2196/42972]
46. Boutilier JJ, Loganathar P, Linden A, Scheer E, Noejovich S, Elliott C, et al. A web-based platform (CareVirtue) to support caregivers of people living with Alzheimer disease and related dementias: mixed methods feasibility study. JMIR Aging 2022 Aug 04;5(3):e36975 [FREE Full text] [doi: 10.2196/36975] [Medline: 35925654]
47. Quinn C, Staub S, Barr E, Gruber-Baldini A. Mobile support for older adults and their caregivers: dyad usability study. JMIR Aging 2019 May 23;2(1):e12276. [doi: 10.2196/12276]
48. Guo Y, Yang F, Hu F, Li W, Ruggiano N, Lee HY. Existing mobile phone apps for self-care management of people with Alzheimer disease and related dementias: systematic analysis. JMIR Aging 2020 Jan 24;3(1):e15290 [FREE Full text] [doi: 10.2196/15290] [Medline: 32012045]
49. Johnson J, Arnold V, Piper AM, Hayes GR. "It's a lonely disease": Cultivating Online Spaces for Social Support among People Living with Dementia and Dementia Caregivers. 2022 Presented at: ACM Conference on Human Factors in Computing Systems; April 30-May 5, 2022; New Orleans, LA. [doi: 10.1145/3555133]
50. Smith CE, Levonian Z, Ma H, Giaquinto R, Lein-Mcdonough G, Li Z, et al. "I Cannot Do All of This Alone". ACM Trans. Comput.-Hum. Interact 2020 Aug 17;27(5):1-41. [doi: 10.1145/3402855]
51. Al-Megren S, Majrashi K, Allwihan RM. ”The Dementia Diva Strikes Again!”: A Thematic Analysis of How Informal Carers of Persons with Dementia Use TikTok. 2021 Presented at: MobileHCI '21: 23rd International Conference on Mobile Human-Computer Interaction; September 27 - October 1, 2021; Toulouse, France. [doi: 10.1145/3447527.3474857]
52. Yuan F, Bowland S, Proctor L, Blackburn J, Mukherjee N, Bray R, et al. Robot-assisted Psycho-education to Enhance Alzheimer’s Caregiver Health. 2022 Presented at: IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE); November 17-19, 2022; Arlington, VA. [doi: 10.1109/chase52844.2021.00027]
53. Lin XY, Saksono H, Stowell E, Lachman ME, Castaneda-Sceppa C, Parker AG. Go and Grow: an evaluation of a pervasive social exergame for caregivers of loved ones with dementia. Proc ACM Hum Comput Interact 2020 Oct;4(CSCW 2):1 [FREE Full text] [doi: 10.1145/3415222] [Medline: 33150303]
54. Granollers T, Lorés J. Incorporation of users in the Evaluation of Usability by Cognitive Walkthrough. In: Navarro-Prieto R, Vidal J, editors. HCI related papers of Interacción 2004. Dordrecht, The Netherlands: Springer; 2006:243-255.
55. Mahatody T, Sagar M, Kolski C. State of the art on the cognitive walkthrough method, its variants and evolutions. International Journal of Human-Computer Interaction 2010 Jul 30;26(8):741-785. [doi: 10.1080/10447311003781409]
56. Lewis C. Using the "Thinking Aloud" Method in Cognitive Interface Design. Yorktown Heights, NY: IBM Thomas J. Watson Research Division; 1982.
57. Rampin R, Rampin V. Taguette: open-source qualitative data analysis. JOSS 2021 Dec;6(68):3522. [doi: 10.21105/joss.03522]
58. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
59. Linders GM, Vaitonytė J, Alimardani M, Mitev KO, Louwerse MM. A realistic, multimodal virtual agent for the healthcare domain. 2022 Presented at: 22nd ACM International Conference on Intelligent Virtual Agents; September 6-9, 2022; Faro, Portugal. [doi: 10.1145/3514197.3551250]

60. Soubutts E, Ayobi A, Eardley R, McNaney R, Cater K, O'Kane AA. Amazon Echo Show as a multimodal human-to-human care support tool within self-isolating older UK households. Proc. ACM Hum.-Comput. Interact 2022 Nov 11;6(CSCW2):1-31. [doi: 10.1145/3555193]
61. Otenen E. Towards Designing for Multimodal Remembering: Findings from an Interview Study. 2024 Presented at: CHI '24: CHI Conference on Human Factors in Computing Systems; May 11-16, 2024; Honolulu, HI URL: https://doi.org/10.1145/3613905.3647965 [doi: 10.1145/3613905.3647965]
62. Kuriakose B, Shrestha R, Sandnes FE. Multimodal navigation systems for users with visual impairments—a review and analysis. MTI 2020 Oct 16;4(4):73. [doi: 10.3390/mti4040073]
63. Hindricks KV, Meyer JJC. Artificial intelligence in health care and medicine: a personalized approach. Acta Scientific Medical Sciences 2019;3(1):71-78. [doi: 10.31080/ASMS.2019.03.0410]
64. Kozyreva A, Lorenz-Spreen P, Hertwig R, Lewandowsky S, Herzog SM. Public attitudes towards algorithmic personalization and use of personal data online: evidence from Germany, Great Britain, and the United States. Humanit Soc Sci Commun 2021 May 14;8(1):1. [doi: 10.1057/s41599-021-00787-w]
65. Seo H, Erba J, Geana M, Lumpkins C. Calling Doctor Google? Technology adoption and health information seeking among low-income African-American older adults. JPIC 2017 Dec 22;1(2):153. [doi: 10.32473/jpic.v1.i2.p153]
66. Santos SLZ, Tagai EK, Scheirer MA, Bowie J, Haider M, Slade J, et al. Adoption, reach, and implementation of a cancer education intervention in African American churches. Implement Sci 2017 Mar 14;12(1):36 [FREE Full text] [doi: 10.1186/s13012-017-0566-z] [Medline: 28292299]
67. Davis CM. Health beliefs and breast cancer screening practices among African American women in California. Int Q Community Health Educ 2020 Jul 13;41(3):259-266. [doi: 10.1177/0272684x20942084]
68. Tagai EK, Scheirer MA, Santos SLZ, Haider M, Bowie J, Slade J, et al. Assessing capacity of faith-based organizations for health promotion activities. Health Promot Pract 2018 Sep 23;19(5):714-723 [FREE Full text] [doi: 10.1177/1524839917737510] [Medline: 29058956]

Abbreviations

ADRD: Alzheimer disease and related dementias
AI: artificial intelligence
GPT: generative pretrained transformer
LLM: large language model
OHC: online health community

Edited by J Wang; submitted 11.06.24; peer-reviewed by M Gasmi , K Bonds Johnson; comments to author 02.10.24; revised version received 15.10.24; accepted 03.12.24; published 08.01.25.

Please cite as:
Bosco C, Otenen E, Osorio Torres J, Nguyen V, Chheda D, Peng X, Jessup NM, Himes AK, Cureton B, Lu Y, Hill CV, Hendrie HC, Barnes PA, Shih PC
Designing a Multimodal and Culturally Relevant Alzheimer Disease and Related Dementia Generative Artificial Intelligence Tool for Black American Informal Caregivers: Cognitive Walk-Through Usability Study
JMIR Aging 2025;8:e60566
URL: https://aging.jmir.org/2025/1/e60566
doi:10.2196/60566
PMID:

Original Paper

Assessment of Technology Readiness in Norwegian Older Adults With Long-Term Health Conditions Receiving Home Care Services: Cross-Sectional Questionnaire Study

Sverre Bergh[1,2], MD, PhD; Jūratė Šaltytė Benth[1,3,4], PhD; Lisbeth Dyrendal Høgset[1], MHS; Britt Rydjord[5], PhD; Lars Kayser[6], Prof Dr Med

[1]Research Centre for Age-related Functional Decline and Disease, Innlandet Hospital Trust, Ottestad, Norway

[2]Norwegian National Centre for Aging and Health, Vestfold Hospital Trust, Tønsberg, Norway

[3]Institute for Clinical Medicine, University of Oslo, Oslo, Norway

[4]Health Services Research Unit, Akershus University Hospital, Oslo, Norway

[5]Department for Research and Innovation, Innlandet Hospital Trust, Ottestad, Norway

[6]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Sverre Bergh, MD, PhD
Research Centre for Age-related Functional Decline and Disease
Innlandet Hospital Trust
Pb 68
Ottestad, 2312
Norway
Phone: 47 45679393
Email: sverre.bergh@sykehuset-innlandet.no

Abstract

Background: With the increasing number of older adults globally, there is a constant search for new ways to organize health care services. Digital health services are promising and may reduce workload and at the same time improve patient well-being. A certain level of eHealth literacy is needed to be able to use digital health services. However, knowledge of technology readiness in this target group of older adults is unclear.

Objective: The aim of this study was to understand the technology readiness level of a group of older adults who were provided home care services in order to address the present and future needs of this group in relation to the implementation of digital health care services.

Methods: This quantitative cross-sectional study included 149 older adults from Norway receiving home care services. The participants completed the Readiness and Enablement Index for Health Technology (READHY) instrument, assessments of well-being (World Health Organization-Five Well-Being Index [WHO-5]), and assessments of demographic and clinical variables (sex, age, education, living situation, comorbidity, use of digital devices, and use of IT). Cluster analyses were used to group the users according to their technology readiness.

Results: The mean participant age was 78.6 (SD 8.0) years, and 55.7% (83/149) were women. There was good consistency within the assumed READHY scales (Cronbach α=.61-.91). The participants were grouped into 4 clusters, which differed in terms of READHY scores, demographic variables, and the use of IT in daily life. Participants in cluster 1 (n=40) had the highest scores on the READHY scales, were younger, had a larger proportion of men, had higher education, and had better access to digital devices and IT. Participants in cluster 4 (n=16) scored the lowest on eHealth literacy knowledge. Participants in cluster 1 had relatively high levels of eHealth literacy knowledge and were expected to benefit from digital health services, while participants in cluster 4 had the lowest level of eHealth literacy and would not easily be able to start using digital health services.

Conclusions: The technology readiness level varied in our cohort of Norwegian participants receiving home care. Not all elderly people have the eHealth literacy to fully benefit from digital health services. Participants in cluster 4 (n=16) had the lowest scores in the eHealth Literacy Questionnaire scales in the READHY instrument and should be offered nondigital services or would need extensive management support. The demographic differences between the 4 clusters may inform stakeholders about which older people need the most training and support to take advantage of digital health care services.

(JMIR Aging 2025;8:e62936) doi:10.2196/62936

KEYWORDS

eHealth literacy; digital health services; technology readiness; Readiness and Enablement Index for Health Technology; READHY; chronic conditions

Introduction

In the coming years, the world will have an increasing number of older adults. According to the World Health Organization (WHO), 1 in 6 people in the world will be 60 years or older in 2030 [1]. Today's aging population is more healthy than previous generations, with a longer life expectancy and more years without frailty [2]. Nevertheless, in the future, there will be more older adults with long-term health conditions and frailty [1]. This will further increase the burden on health care services in the coming years. This change in demographics together with a reduced workforce due to a reduction in the birth rate calls for new ways to organize health and care services. There is an urgent need for more efficient ways to provide the needed health care with fewer in-person contacts, the involvement of informal caregivers, an increase in self-management, and a reduction in travel time, particularly in rural areas. Increased digitalization of health care services could be a solution, which could enable and engage more patients, resulting in increased levels of self-efficacy, self-management, and empowerment [3], and thus reduce the burden on the workforce [4]. Moreover, telehealth solutions may increase the number of virtual contacts and provide means for self-monitoring and dialogues with health professionals as an immediate response to the early signs of deterioration in the health condition [5]. The use of telemedicine may be of particular importance in rural areas with scattered settlements, saving travel time and costs for both health personnel and patients. Although evidence for the assumed benefits of the digital transformation of health care services is scarce, particularly among older adults, it will continue in the coming years along with the general digitalization of both private and public services in Norway and many other countries.

In 2021, 96% of Norwegians older than 9 years had access to a smartphone [6]. However, this indicates that 4%, or approximately 200,000 people, do not have a smartphone. A higher number of people probably have an insufficient level of digital health literacy. In a study from 2021, the Norwegian Directorate of Health found that 13% of people aged ≥60 years never used the internet, while 9% used the internet only once a week. Moreover, only 66% of people aged ≥80 years used the internet. Further, 44% of people aged ≥60 years reported no or low experience with digital IDs and 2-factor identification, which are needed to access public web pages [7].

One benefit from the digital health transformation is that citizens and patients may communicate from their homes and receive care and treatment virtually. Services may also be provided by specialists at hospitals in collaboration with registered nurses in municipalities and the person's own general practitioner, offering better coordination. This requires the individuals being cared for to have access to the internet; have devices, such as smartphones, tablets, laptops, and PCs; and have the ability to use these devices. For those not having these resources, it is essential to identify this challenge and plan for the involvement of supporting resources or nondigital services. This calls for new ways to plan for digitally enabled services and usage of digital tools, and to be aware of the specific resources and needs of individuals when offering these.

In this context, awareness of the person's level of digital health literacy [8] and level of technology readiness [9] is important and may be helpful information together with sociodemographic characteristics, including access to and familiarity with digital devices. In particular, older adults have a lower level of digital health literacy, and as lower levels of digital health literacy may be related to economic status, it may be hypothesized that older people living in rural areas with precisely these characteristics may have a low or even insufficient level of digital health literacy. Technology readiness may be a barrier for offering telehealth solutions, which otherwise would be of huge benefit to these people, due to the fact that the remote areas of living may prohibit access to in-person services, particularly outside the summer season. This calls for knowledge about the levels of digital health literacy and technology readiness of different populations.

eHealth literacy, which is currently often used synonymously with digital health literacy, was defined in 2006 by Norman and Skinner as "the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem" [8]. Several factors, such as frailty, low level of digital health literacy, and motoric or cognitive impairment, may be barriers for patients when introducing new technologies [10].

One way to identify barriers is by measuring the level of digital health literacy using a multi-faceted instrument such as the eHealth Literacy Questionnaire (eHLQ) [11]. Another way is to expand on the factors necessary for being able to apply digital health literacy in a context, by using the Readiness and Enablement Index for Health Technology (READHY) [9]. The READHY instrument builds on the eHLQ but includes 4 scales related to self-management from the Health Education Impact Questionnaire (heiQ) [12] and 2 scales related to support from the Health Literacy Questionnaire (HLQ) [13]. The READHY instrument has been proven to inform about variations in needs and skills in relation to technology readiness presented as specific clusters.

The READHY instrument can be used both as a framework for qualitative studies [14] and as an instrument to assess the experience of support, elements of self-management, and levels of digital health literacy. It has been used in a number of contexts, such as cancer survivors [15], type 2 diabetes mellitus (T2DM) patients [16], older hospitalized medical patients, and people living with inflammatory bowel disease [17,18]. The separate scales of the READHY instrument have been translated and validated in the Norwegian language, but the complete

READHY instrument has not yet been published in Norwegian research [19,20].

Despite emerging knowledge within various specific conditions, studies focusing on the technology readiness of older adults are lacking. People receiving home care services may have other needs for digital health services than people with specific diseases, such as cancer and T2DM, and it is important to have an instrument that is broad enough to catch these differences. In our study, the READHY instrument will be tested as an option to elucidate the characteristics of this population of older adults living in a rural area.

Data obtained using the READHY instrument can be clustered to identify segments of users, and this segmentation can be enriched with other characteristics, such as sociodemographics, health conditions, and well-being, forming clusters to help health professionals recognize the need for support or training among the people they serve [21]. This information may help health service organizations to plan for subpopulations based on the identified clusters of the general population and with knowledge about the size of particular groups. It may also help health care professionals to have better conversations when introducing new digital health tools or offering telehealth services to ensure meaning and motivation for older adults [22].

The aim of this study was to explore the technology readiness level in a population of older adults living in a rural area, a group that is hypothesized to be less technology ready. Further, the study attempted to examine their digital resources and behaviors and evaluate associations between these factors and their self-reported health and mental well-being, which may influence their motivation and ability to create meaning in using digital health services and tools.

The results of the study may help service providers to better plan for the specific needs of subgroups and provide health care professionals with insights that may help them to better understand how to introduce new digital health tools and services in this group of older adults living in rural areas. This group is more likely to be frail and at the same time may benefit more from virtual and other contacts that are not in-person. Using clustering analysis, we explored the technology readiness level of older adults receiving home care and their characteristics in terms of resources and digital behaviors (ie, sociodemographic variables and use of digital devices).

Methods

Design

This cross-sectional study included 149 older adults recruited from 7 municipalities in South-Eastern Norway. The data were collected by 1 to 2 health care staff in each of the 7 participating municipalities. The same research nurse trained all data collectors (2-hour training session), and the data collectors were supervised by the research nurse during the study period. The participants completed the questionnaires listed below, either self-reported by themselves or in an interview with the health care staff. The participants were approached in a face-to-face setting where they were informed about the study, signed the consent form, and completed the questionnaires.

Population

Health care staff in the municipal home care services and at the lung department of Innlandet Hospital Trust recruited the participants. Patients admitted to the outpatient clinic at the lung department and people receiving home care services from the municipalities were invited to participate in the study. The inclusion criteria were as follows: (1) living at home, (2) ≥65 years of age, (3) receiving home care services, (4) presence of long-term health conditions, (5) fluency in Norwegian, and (6) written informed consent. The exclusion criteria were as follows: (1) severe dementia (Clinical Dementia Rating scale score of 3), (2) terminal disease with <3 months of expected survival, and (3) severe psychiatric disease, including drug abuse. A total of 211 individuals were eligible, of which 62 were not included as they did not consent to participate. Those not included were on average 83.5 (SD 5.7) years old, and 61.3% were female. The process of participant inclusion was from September 2021 to June 2023.

Included Data and Questionnaires

We collected data on age (years), sex (male/female), living situation (alone, with spouse/partner, or with others), comorbidity, education (compulsory lower secondary school, upper secondary education, master's degree, or more than 4 years university), use of digital devices (laptop, PC, tablet, or smartphone), use of technology to access the general physician and electronic medical record, and digital competence (very poor, poor, average, good, or very good).

The READHY instrument consists of 65 items, with 4 scales from the heiQ [12], 2 scales from the HLQ [13], and all scales from the eHLQ (Table 1) [11]. The heiQ is used to assess patient education and self-management and is validated in Norwegian [12,19]. The HLQ is used to assess health literacy and has been translated to and validated in many languages including Norwegian [13,23]. The eHLQ is used to assess the benefits of telemedicine for individuals, evaluate how telehealth services work together, and examine the telemedicine knowledge of individuals [18]. It has been translated, validated, and psychometrically tested in Norwegian [24]. The full READHY instrument is validated in several languages but not yet in Norwegian. The READHY instrument assesses the person's health technology readiness [9]. Each item in the READHY instrument can be scored from 1 to 4 (strongly disagree, disagree, agree, and strongly agree). The scoring of "strongly agree" in the item heiQ8 *emotional distress* is reversed for the statistical analysis, with the highest score representing the lowest level of distress. The scores of the scales in the READHY instrument are calculated as the mean scores of the included items for each scale. The scores of scales with missing data have been included if the number of missing items is ≤50%.

Table 1. Scales from the Health Education Impact Questionnaire (heiQ), Health Literacy Questionnaire (HLQ), and eHealth Literacy Questionnaire (eHLQ) included in the Readiness and Enablement Index for Health Technology (READHY).

Questionnaire and scale	Description
heiQ[a]	
heiQ3	Self-monitoring and insight
heiQ4	Constructive attitudes and approaches
heiQ5	Skill and technique acquisition
heiQ8	Emotional distress
HLQ[b]	
HLQ1	Feeling understood and supported by health care providers
HLQ4	Social support for health
eHLQ[c]	
eHLQ1	Using technology to process health information
eHLQ2	Understanding of health concepts and language
eHLQ3	Ability to actively engage with digital services
eHLQ4	Feeling safe and in control
eHLQ5	Motivated to engage with digital services
eHLQ6	Access to digital services that work
eHLQ7	Digital services that suit individual needs

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

Cutoff values are not meaningful from a psychometric approach, and the values for scoring in the READHY scales have not been calibrated to each other or to outcomes. Despite this, scores below 2.7 on the scales may mirror that the person has a level of "not being sufficient" as not all items in the scales are scored as "strongly agree" or "agree." A score of 2.0 or lower on the scales may be considered problematic for managing digital health services as most items in the scales are scored as "strongly disagree" or "disagree" [25]. The READHY instrument takes 10 to 20 minutes to complete in most populations (including older adults), but it could take up to 30 minutes to complete for individuals with cognitive impairment.

The World Health Organization-Five Well-Being Index (WHO-5) consists of 5 items. Each item is scored from 0 to 5, where 5 indicates *all the time* and 0 indicates *at no time*. The score of each item is multiplied by 4, giving a total score from 0 to 100.

Some additional questions were asked to enrich the data. A question about digital competence was asked: "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" A question about the participants' health was taken from the 36-item Short-Form Survey instrument (SF-36): "In general, would you say your health is excellent/very good/good/fair/poor?" Scores range from 1 (excellent) to 5 (poor) [26], and they were treated as a continuous variable.

Statistical Analysis

The participant characteristics have been presented as means and SDs or frequencies and percentages, as appropriate. The READHY scale score was calculated as the mean of the item scores when fewer than half of the items had missing values. This resulted in 2 excluded cases where the scale score could not be calculated. Ideally, one should start with assessing the dimensionality of the READHY instrument by applying a factor analysis. However, due to convergence problems caused by a very small sample size with respect to the number of items, this approach was not possible. We therefore report the Cronbach α for the scales identified by a previous factor analysis of a Danish dataset [9].

To identify groups of participants based on the READHY scales, we first applied a hierarchical cluster analysis with Euclidean squared distance and Ward linkage to identify the possible number of groups [27]. The results of the hierarchical cluster analysis were then used to determine an appropriate number of clusters in the next step, a k-means cluster analysis [28] with Euclidean similarity measure. By examining the dendrogram from the first step, we expected to gain insights into the natural grouping of data, which would inform the subsequent k-means analysis. In this way, we leveraged the strengths of both methods for more effective clustering and based our conclusion on k-means analysis, which is easier to interpret due to each individual being assigned to one cluster. While choosing the final number of clusters, we considered clinical relevance,

cluster size, small within-cluster variability, and between-cluster heterogeneity.

One-way ANOVA was then applied to assess the differences between the identified clusters with respect to the READHY scales. Further, one-way ANOVA, the chi-square test, and the Fisher exact test were applied to compare the profiles of the identified clusters based on the following variables: all scale scores, age, sex, education (≤12 years or >12 years), living situation (alone, with spouse/partner, or with others), comorbidity (number of chronic diagnoses), use of a laptop, PC, or tablet at least once a week, use of a smartphone at least once a week, use of IT in previous work or studies, use of IT to communicate with public services, use of the Norwegian digital identifier to access a webpage, use of IT to communicate with a general practitioner, logging into a national health webpage, use of IT to find health information on the internet or social media, reading medical journals or test results on a national health webpage, IT competence assessed by others (very poor, poor, average, good, or very good), WHO-5 score, and response to the question "In general, would you say your health is excellent/very good/good/fair/poor?" Pairwise comparisons of groups were performed as post hoc analyses. The results of these post hoc analyses were reported as mean differences and SDs for continuous variables, mean differences and SDs in proportions for dichotomous variables, and effect sizes calculated as absolute values of z-statistics from the Mann-Whitney *U* test divided by the square root of the total sample size for ordinal variables, along with *P* values. Only significant pairwise differences were tabulated.

Two linear regression models were estimated to assess the associations of WHO-5 (primary) and IT competence assessed by others (secondary) with prechosen covariates: age, sex, education (≤12 years or >12 years), living situation (alone, with spouse/partner, or with others), comorbidity, use of a laptop, PC, or tablet at least once a week, use of a smartphone at least once a week, use of IT to communicate with a general practitioner, use of IT to find health information on the internet or social media, reading medical journals or test results on a national health webpage, and response to the question "In general, would you say your health is excellent/very good/good/fair/poor?" The regression models were estimated in cases with no missing values for covariates.

Statistical analyses were performed using STATA version 18 (StataCorp). A *P* value of <.05 was considered statistically significant. No adjustment for multiple testing was applied; however, we separated between ANOVA as the main analysis and post hoc analyses considered purely exploratory. Although we have only reported the results of the post hoc analyses with *P*<.05, we emphasize that all pairwise comparisons were performed. Further, the regression analyses have been considered secondary. Moreover, there is clarity regarding the number of tests performed, allowing the reader to consider the reliability of our findings.

Ethical Considerations

Participation in the study was based on informed written consent. The study was considered by The Regional Committee for Medical and Health Research Ethics in Norway as being outside its jurisdiction. The data collection, data storage, data analysis, and publication of the results were approved by the data protection officer at Innlandet Hospital Trust (number: 14832226). Participants could withdraw from the study at any time and request for their data to be deleted.

Results

The descriptive results of the cohort are presented in Table 2. The mean age of the participants was 78.6 (SD 8.0) years, and 55.7% (83/149) were women. Moreover, 19.2% (27/141) had higher education (>12 years) and 69.8% (104/149) were living alone. The mean scores of the READHY scales ranged from 2.0 (SD 0.5) for the eHLQ1 to 3.1 (SD 0.3) for the heiQ3. The heiQ scales assessed patient education and self-management, and the mean scores for the 4 scales ranged from 2.8 (SD 0.6) to 3.1 (SD 0.3), indicating high levels of patient education and self-management. The HLQ scales assessed health literacy, and the mean scores ranged from 2.9 (SD 0.4) to 3.0 (SD 0.5), indicating a high level of health literacy. The eHLQ scales assessed eHealth literacy, and the mean scores ranged from 2.0 (SD 0.5) to 3.0 (SD 0.3), indicating that eHealth literacy was low in our population. Moreover, the scores differed between the scales. Table 3 presents Cronbach α values for the READHY scales. Except for the eHLQ2 scale with a low Cronbach α of .61, the Cronbach α values ranged from .70 to .91 for the READHY scales, showing good consistency within the assumed scales [29].

Table 2. Summarized statistics of the cohort.

Variable	Value (N=149)
Age (years) (n=148), mean (SD)	78.6 (8.0)
Gender (female), n (%)	83 (55.7)
Education (n=141), n (%)	
≤12 years	114 (80.9)
>12 years	27 (19.2)
Living situation, n (%)	
Living alone	104 (69.8)
Living with spouse/partner/others	45 (30.2)
Number of chronic diagnoses (n=144), mean (SD)	2.6 (1.4)
Use of digital devices at least once a week, n (%)	
PC, laptop, or tablet	72 (48.3)
Smartphone	89 (59.7)
IT to communicate with a general practitioner	34 (22.8)
IT to find health information on the internet or social media	17 (11.6)
Reading medical journals or test results on a national health webpage, n (%)	
Never/rarely	116 (78.4)
Sometimes/always	32 (21.6)
In general, would you say your health is excellent/very good/good/fair/poor, mean (SD)	1.2 (0.8)
READHY[a] scales, mean (SD)	
heiQ[b]	
heiQ3	3.1 (0.3)
heiQ4	2.9 (0.4)
heiQ5	2.8 (0.4)
heiQ8	2.8 (0.6)
HLQ[c]	
HLQ1 (n=148)	3.0 (0.5)
HLQ4 (n=148)	2.9 (0.4)
eHLQ[d]	
eHLQ1 (n=147)	2.0 (0.5)
eHLQ2	2.8 (0.4)
eHLQ3 (n=147)	2.2 (0.6)
eHLQ4	3.0 (0.3)
eHLQ5	2.2 (0.5)
eHLQ6	2.4 (0.4)
eHLQ7 (n=148)	2.2 (0.6)

[a]READHY: Readiness and Enablement Index for Health Technology.

[b]heiQ: Health Education Impact Questionnaire.

[c]HLQ: Health Literacy Questionnaire.

[d]eHLQ: eHealth Literacy Questionnaire.

Table 3. Cronbach α values of the Readiness and Enablement Index for Health Technology (READHY) scales.

Scale	Cronbach α
heiQ[a]	
heiQ3: Self-monitoring and insight (6 items)	.69
heiQ4: Constructive attitudes and approaches (5 items)	.77
heiQ5: Skills and technique acquisition (4 items)	.70
heiQ8: Emotional distress (6 items)	.90
HLQ [b]	
HLQ1: Feeling understood and supported by health care providers (4 items)	.83
HLQ4: Social support for health (5 items)	.77
eHLQ[c]	
eHLQ1: Using technology to process health information (5 items)	.86
eHLQ2: Understanding of health concepts and language (5 items)	.61
eHLQ3: Ability to actively engage with digital services (5 items)	.91
eHLQ4: Feeling safe and in control (5 items)	.71
eHLQ5: Motivated to engage with digital services (5 items)	.80
eHLQ6: Access to digital services that work (6 items)	.78
eHLQ7: Digital services that suit individual needs (4 items)	.85

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

Based on visually inspecting a dendrogram from the hierarchical cluster analysis (not presented), exploring 3 to 6 cluster solutions by k-means cluster analysis, and considering clinical relevance, within-cluster variability, between-cluster heterogeneity, and reasonable sample size, the 4-cluster model was chosen (Table 4). The details of patient profiles in the 4 clusters are described in Multimedia Appendix 1. The participants in cluster 1 (n=40) had on average the highest scores on the READHY scales (meaning better eHealth literacy), except the HLQ4 scale. The HLQ4 scale assesses the social support for health, where the scores were intermediate. Participants in cluster 1 were younger than the other participants (P=.006), had a larger proportion of men (P<.001), had higher education (P<.001), had better access to smartphones and PCs, laptops, or tablets (P<.001), had used IT more in their previous work (P<.001), and currently used IT to access health information and stay in touch with the health care system (P<.001). The participants in cluster 3 (n=25) had the lowest average scores on the heiQ4, heiQ5, and heiQ8 scales (patient education and self-management) and the 2 HLQ scales (feeling understood and supported by health care providers, and social support for health). The participants in cluster 4 (n=16) had the lowest average scores on the eHLQ scales (meaning low eHealth literacy knowledge), except the eHLQ4 scale (feeling safe and in control), where their scores were comparable to those of the 3 other clusters. Thus, the feelings of being safe and in control were equal in the 4 clusters. Cluster 2 was the largest (n=66) and had intermediate scores on most of the scales. Interestingly, the question “IT competence assessed by others” differentiated between all clusters, except between cluster 2 and cluster 3 (P<.001).

Table 4. Results of k-means clustering with 4 clusters (n=147).

Scale	Cluster 1 (n=40), mean (SD)	Cluster 2 (n=66), mean (SD)	Cluster 3 (n=25), mean (SD)	Cluster 4 (n=16), mean (SD)	*P* value for ANOVA	*P* value for pairwise comparisons, mean difference (SD) and *P* value
heiQ[a]						
heiQ3	3.2 (0.3)	2.9 (0.3)	3.0 (0.3)	3.2 (0.5)	<.001	1 vs 2: 0.3 (0.1), *P*<.001; 1 vs 3: 0.2 (0.1), *P*=.01; 2 vs 4: –0.2 (0.1), *P*=.009
heiQ4	3.1 (0.3)	3.0 (0.3)	2.4 (0.3)	3.0 (0.5)	<.001	1 vs 3: 0.7 (0.1), *P*<.001; 2 vs 3: 0.6 (0.1), *P*<.001; 3 vs 4: –0.6 (0.1), *P*<.001
heiQ5	2.9 (0.4)	2.8 (0.3)	2.6 (0.5)	3.0 (0.5)	.001	1 vs 2: 0.1 (0.1), *P*=.047; 1 vs 3: 0.4 (0.1), *P*=.001; 2 vs 3: 0.3 (0.1), *P*=.004; 3 vs 4: –0.4 (0.2), *P*=.01
heiQ8	2.9 (0.6)	3.0 (0.3)	2.1 (0.5)	2.8 (0.7)	<.001	1 vs 3: 0.8 (0.1), *P*<.001; 2 vs 3: 0.9 (0.1), *P*<.001; 3 vs 4: –0.7 (0.2), *P*=.002
HLQ[b]						
HLQ1	3.2 (0.5)	2.9 (0.4)	2.9 (0.3)	3.0 (0.7)	.003	1 vs 2: 0.3 (0.1), *P*<.001; 1 vs 3: 0.3 (0.1), *P*=.006
HLQ4	3.0 (0.4)	2.8 (0.4)	2.6 (0.4)	3.2 (0.4)	<.001	1 vs 3: 0.3 (0.1), *P*=.004; 2 vs 3: 0.2 (0.1), *P*=.009; 2 vs 4: –0.4 (0.1), *P*=.001; 3 vs 4: –0.6 (0.1), *P*<.001
eHLQ[c]						
eHLQ1	2.5 (0.5)	2.0 (0.2)	2.1 (0.3)	1.1 (0.2)	<.001	1 vs 2: 0.6 (0.1), *P*<.001; 1 vs 3: 0.4 (0.1), *P*=.001; 1 vs 4: 1.4 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.002; 2 vs 4: 0.8 (0.1), *P*<.001; 3 vs 4: 1.0 (0.1), *P*<.001
eHLQ2	3.1 (0.3)	2.6 (0.3)	2.8 (0.2)	2.6 (0.5)	<.001	1 vs 2: 0.4 (0.1), *P*<.001; 1 vs 3: 0.3 (0.1), *P*<.001; 1 vs 4: 0.4 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.02
eHLQ3	2.9 (0.5)	2.1 (0.3)	2.3 (0.4)	1.1 (0.2)	<.001	1 vs 2: 0.8 (0.1), *P*<.001; 1 vs 3: 0.6 (0.1), *P*<.001; 1 vs 4: 1.8 (0.1), *P*<.001; 2 vs 3: –0.2 (0.1), *P*=.004; 2 vs 4: 1.0 (0.1), *P*<.001; 3 vs 4: 1.2 (0.1), *P*<.001
eHLQ4	3.1 (0.4)	3.0 (0.2)	3.0 (0.2)	3.1 (0.4)	.08	1 vs 2: 0.1 (0.1), *P*=.02
eHLQ5	2.7 (0.5)	2.0 (0.2)	2.1 (0.3)	1.5 (0.4)	<.001	1 vs 2: 0.6 (0.1), *P*<.001; 1 vs 3: 0.6 (0.1), *P*<.001; 1 vs 4: 1.2 (0.1), *P*<.001; 2 vs 4: 0.5 (0.1), *P*<.001; 3 vs 4: 0.6 (0.1), *P*<.001
eHLQ6	2.9 (0.3)	2.3 (0.2)	2.3 (0.3)	1.6 (0.3)	<.001	1 vs 2: 0.6 (0.05), *P*<.001; 1 vs 3: 0.5 (0.1), *P*<.001; 1 vs 4: 1.2 (0.1), *P*<.001; 2 vs 4: 0.6 (0.1), *P*<.001; 3 vs 4: 0.7 (0.1), *P*<.001
eHLQ7	2.8 (0.4)	2.0 (0.2)	2.0 (0.3)	1.2 (0.4)	<.001	1 vs 2: 0.8 (0.1), *P*<.001; 1 vs 3: 0.8 (0.1), *P*<.001; 1 vs 4: 1.6 (0.1), *P*<.001; 2 vs 4: 0.8 (0.1), *P*<.001; 3 vs 4: 0.8 (0.1), *P*<.001

[a]heiQ: Health Education Impact Questionnaire.

[b]HLQ: Health Literacy Questionnaire.

[c]eHLQ: eHealth Literacy Questionnaire.

In the bivariate linear regression model, WHO-5 as an outcome (Multimedia Appendix 2) was significantly associated with education and the question “In general, would you say your health is excellent/very good/good/fair/poor?”, while only the latter question remained significant in the multiple model. A 1-point increase in the score of the question “In general, would you say your health is excellent/very good/good/fair/poor?” was associated with a 12.7-point increase in the WHO-5 score.

In the multiple linear regression model with the question “IT competence assessed by others” as an outcome (Multimedia Appendix 3), only “use of a PC, laptop, or tablet at least once in the last week” and “reading medical journals or test results on a national health webpage” were significantly associated with the outcome.

Discussion

Summary of the Main Findings

In this study, we included participants from a rural area in Norway with less education than other parts of Norway, and our recruitment procedure resulted in the inclusion of slightly more women than men. The participants were assessed for technology readiness using the READHY instrument. We found that people receiving home care services in this area had high

levels of patient education and self-management (scores of 2.8-3.1) and high levels of health literacy (scores of 2.9-3.0) but generally low levels of eHealth literacy (scores of 2.0-3.0). The findings were based on the assumption that a score of ≥2.7 (out of a maximum of 4) reflects a sufficient level, as this score reflects overweight of the response "agree" or "strongly agree" to all the items in the scale [25]. The level of eHealth literacy also differed, as scores on the eHLQ scales and between the participants differed, showing less stability than the heiQ and HLQ scales. The eHealth scores were lower in this study than in a previous study by Kayser et al [11], which can be explained by differences in age, education, and study settings. Therefore, in a rural population like the population in our study, the low level of eHealth literacy should be considered when introducing digital health services in the home care setting.

Low eHealth Literacy

The most interesting results for health care services were the characteristics of the participants in cluster 4, with the lowest scores on the eHLQ scales, indicating low eHealth literacy knowledge (Multimedia Appendix 1). These participants were older, were mainly women, and had a lower level of education than participants in cluster 1 (with the highest level of eHealth literacy). Participants in cluster 4 seldom used a PC, laptop, tablet, or smartphone, and they seldom accessed public services generally and health care services specifically compared with participants in cluster 1. The participants in cluster 4 had several diagnoses, although not significantly more than that among participants in cluster 1. We expect their need for care and support from the home care service to be high, and this population group could benefit from digital home monitoring. Nevertheless, the results show that these participants would need more help and facilitation from the health care service than others to take advantage of digital health services. Offering nontechnological services may be a better option. Social resources that could support them in the use of digital health services must be identified if digital health services are offered. If not, they could be discriminated from receiving the same health care services as others.

Screening for eHealth Literacy

We found that the response to the question "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" was associated with several variables in the bivariate model, but only the responses to the items "use of PC, laptop, or tablet at least once a week" and "reading medical journals or test results on a national health webpage" remained significant in the multiple model (Multimedia Appendix 3). Surprisingly, no other variables remained significant in the multiple model, but the most likely explanation was the low power in our study. Further, we found a strong association between belonging to 1 of the 4 clusters and the question "If someone knowing you well should assess your IT competence, would they say it is very poor/poor/average/good/very good?" (Multimedia Appendix 1). The answer differentiated between participants in cluster 1 and those in clusters 2, 3, and 4; between participants in cluster 2 and those in cluster 3; and between participants in cluster 3 and those in cluster 4. These associations contribute to the increasing amount of data supporting the relevance and content validity of this single-item scale and indicate that a single question could be used to screen larger populations for the level of digital health skills. Those scoring low on the single question could be further assessed with multidimensional instruments such as the READHY instrument. The READHY instrument consists of 13 scales with a total of 65 items and takes 10 to 20 minutes to complete. It may be exhausting to complete, especially for old frail people with cognitive impairment. Thus, it is preferable if clinicians can scan a population and collect the same information about the technological readiness level using a single question. This single question could potentially help to increase the number of screenings to identify those in need of further investigation. Here, the READHY instrument can be used to identify specific resources and challenges in relation to the use of digitally enabling or assisting services and technologies. However, it is important to emphasize that using 1 screening item will lead to a loss in the understanding of where particular challenges are. Further, we should be aware that it only screens for digital expertise, and self-management and social support are not taken into consideration with this approach.

Well-Being

We analyzed the variables associated with the scores of the WHO-5 (Multimedia Appendix 2). In the bivariate model, there was an association between a higher educational level and higher scores on the WHO-5, which was not significant when adjusting for other variables such as age, sex, living situation, and number of chronic diagnoses. People with a higher level of education had better health, were younger, and had better digital health literacy, which contributed to the loss of the bivariate effect of education on well-being (WHO-5) when adjusting for these variables. Further, we did not find associations of the co-habitant status and number of chronic diagnoses with well-being assessed by the WHO-5. These findings need some consideration, as other European studies have found that living alone is associated with a lower score on the WHO-5 [30-32]. One explanation may be that the context in these studies differs from that in our study (cohort living in a rural area and having long travel distances). The participants in our study were already receiving home care services and thus were in regular contact with a support person and had access to a service they could call, which may be of importance for well-being. The home care service may be equivalent to the effect of living together with others, as people receiving home care services are in regular contact with a support person and have access to a service they can call.

Cluster Profiles

Our analysis resulted in 4 cluster profiles based on the k-means clustering analysis (Table 4). One cluster with high scores in all the READHY scales, 1 with low scores in the eHLQ scales (eHealth literacy), 1 with low scores in the heiQ scales (health education) and the HLQ scales (health literacy), and 1 with intermediate scores. Being male, being younger, having higher education, and having good access to digital devices and knowledge of IT were all associated with a better technology readiness level. This result was comparable to that in a study by Kayser et al [9], which had 1 cluster with the highest scores

in the READHY scales, 1 with low scores in the eHLQ scales, and 2 involving intermediate profiles. In the Danish study by Kayser et al [9], 25% of the participants were in the cluster with the highest READHY scale scores, which is similar to the proportion of 27% in our study. When considering another study [11], comparable proportions of participants were in the group with the lowest eHLQ scale scores (13% in the previous study and 11% in our study). In another Danish study by Thorsen et al [16], it was found that 5 clusters were more meaningful for the population of patients with T2DM. They also described 1 cluster with the highest READHY scale scores and 1 with low scores on the eHLQ scales, with the 3 remaining clusters having intermediate scores [16]. In a study with Danish cancer survivors, individuals in the cluster with the highest READHY scale scores were younger than those in the other clusters [9], as in our study. On the other hand, in the study of Danish patients with T2DM, 1 of the intermediate clusters had the lowest age [16]. A recent Norwegian study recruiting 260 patients from medical and surgical wards found 6 HLQ clusters and 6 eHLQ clusters [33]. They did not collect heiQ data but found that the participants with the lowest HLQ and eHLQ scores had comparable socioeconomic clusters, similar to the participants in cluster 4 in our study [33]. Thus, the 4 profiles based on the k-means clustering analysis of our data are comparable with those in 2 Danish studies and 1 Norwegian study, despite the mean age being higher in our study.

As explained in the statistical analysis subsection in the Methods section, a confirmatory factor analysis to assess the psychometric properties of the READHY instrument was considered. However, it resulted in convergence problems, most likely due to a very small sample size with respect to the number of items. Therefore, we did not succeed in our aim to validate the Norwegian version of the READHY instrument. Instead, we reported the Cronbach α for the scales identified by a previous factor analysis of a Danish dataset [9], which showed good consistency within the assumed scales.

Limitations

The study had several limitations. No formal power calculation was performed prior to the study, and a convenience sample was used. It is suggested that cluster analysis should only be applied when a good group separation is expected, and it is recommended to include a sample size of approximately 20 to 30 individuals per expected group [34]. Based on previous cluster analysis of the READHY instrument [9,16,33], we expected 3 to 5 clusters of individuals in our study. Thus, a sample size of 149 was sufficient according to recommendations. The inclusion of more participants in the study could strengthen our results, could allow us to perform a factor analysis, and might make it possible to report more associations from the regression models (Multimedia Appendix 2 and 3).

The data were collected by 1 to 2 health care staff in each of the 7 participating municipalities. The same research nurse trained all data collectors, and all had experience in assessing and testing the patients. However, we cannot rule out that the use of several data collectors may have caused biases in the data. On the other hand, we used well-established assessment tools that have been found to be reliable and valid. We used a long and extensive case report form (CRF) with several questionnaires. Some of the questionnaires included many questions. The participants may have become tired, especially those with impaired cognitive function, which might have influenced the data quality. The participants were given the opportunity to answer the questionnaires at their own pace and could take breaks if needed. In the end, we found that the data collected were of excellent quality, with limited missing data.

We recruited participants from only 7 municipalities in a rural region in South-Eastern Norway, with a lower education level than that generally in Norway. This fact, in addition to the fact that the participants were recruited from among those receiving home care services, may reduce the generalizability of the study. The READHY instrument should be tested in home care settings in other regions of Norway and in other countries to increase generalizability. The results from future studies and our study could be extrapolated to inform stakeholders in home care services about the need for support when introducing digital health services in their populations. The results may not be generalizable to countries without home care services or with other access to digital devices and IT. On the other hand, the participants in the study represent people who are particularly in need of digital home monitoring because of the travel distances in rural areas, and the results may be important for stakeholders and leaders of health care services in Norway and internationally. Moreover, 62 people did not consent to participate. These people were older and a larger proportion were women when compared with the cohort of participants.

Conclusion

To our knowledge, this is the first study on eHealth literacy in people receiving home care, a population that currently is being offered and in the future will be offered digital technology. Understanding the eHealth literacy knowledge of this population is important to take advantage of digital technology. The technology readiness level varied in the cohort of Norwegian participants receiving home care, and not all people receiving home care services can be expected to benefit from digital health services. Participants in cluster 4 (n=16) had the lowest score in the eHLQ scales in the READHY instrument and should be offered nondigital services or would need extensive management support. The demographic differences between the 4 clusters may inform stakeholders about which older people need the most training and support to be able to take advantage of digital health care services. This study contributes to this important area with information about technology readiness in a population of older adults and with insights into how this is related to their perceived health and digital behaviors. We were not able to report the psychometric properties of the Norwegian version of the READHY instrument, and a larger validation study should be performed.

Acknowledgments

This research was funded by the European Union's Horizon 2020 Research and Innovation Program under grant 101016848 and the Canadian Institutes of Health Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Patient profiles within the 4 clusters (N=147).
[DOCX File , 21 KB - aging_v8i1e62936_app1.docx]

Multimedia Appendix 2
Results of linear regression analysis for the World Health Organization-Five Well-Being Index (WHO-5) (N=129).
[DOCX File , 39 KB - aging_v8i1e62936_app2.docx]

Multimedia Appendix 3
Results of linear regression analysis for the question "IT competence assessed by others" (N=131).
[DOCX File , 39 KB - aging_v8i1e62936_app3.docx]

References

1. Decade of Healthy Ageing. World Health Organization. URL: https://apps.who.int/gb/ebwha/pdf_files/EB146/B146_23-en.pdf [accessed 2025-01-20]
2. Chang M, Michelet M, Skirbekk V, Langballe EM, Hopstock LA, Sund ER, et al. Trends in the use of home care services among Norwegians 70+ and projections towards 2050: The HUNT study 1995-2017. Scand J Caring Sci 2023 Sep;37(3):752-765. [doi: 10.1111/scs.13158] [Medline: 36967552]
3. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 Dec;2:100171 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
4. Stoumpos AI, Kitsios F, Talias MA. Digital transformation in healthcare: technology acceptance and its applications. Int J Environ Res Public Health 2023 Feb 15;20(4):3407 [FREE Full text] [doi: 10.3390/ijerph20043407] [Medline: 36834105]
5. Hong P, Chen K, Chang Y, Cheng S, Chiang H. Effectiveness of theory-based health information technology interventions on coronary artery disease self-management behavior: a clinical randomized waitlist-controlled trial. J Nurs Scholarsh 2021 Jul;53(4):418-427 [FREE Full text] [doi: 10.1111/jnu.12661] [Medline: 33844425]
6. StatBank. Statistics Norway. URL: https://www.ssb.no/en/ [accessed 2025-01-20]
7. Bjønnes A, Midtbø T, Størset H, Ulven C. Befolkningens digitale kompetanse og deltakelse - Med et ekstra blikk på seniorer og ikke-sysselsatte. Direktoratet for høyere utdanning og kompetanse. 2021. URL: https://hkdir.no/rapporter-undersokelser-og-statistikk/befolkningens-digitale-kompetanse-og-deltakelse-2021 [accessed 2025-01-20]
8. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
9. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional readiness and enablement index for health technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
10. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: A systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
11. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen KB, Elsworth G, et al. A multidimensional tool based on the eHealth literacy framework: development and initial validity testing of the eHealth Literacy Questionnaire (eHLQ). J Med Internet Res 2018 Feb 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]
12. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
13. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013 Jul 16;13:658 [FREE Full text] [doi: 10.1186/1471-2458-13-658] [Medline: 23855504]
14. Kikkenborg J, Magelund E, Riise MS, Kayser L, Terp R. Knowledge, skills, and experience with technology in relation to nutritional intake and physical activity among older adults at risk of falls: semistructured interview study. JMIR Hum Factors 2024 May 08;11:e52575 [FREE Full text] [doi: 10.2196/52575] [Medline: 38717810]

15. Rossen S, Kayser L, Vibe-Petersen J, Ried-Larsen M, Christensen JF. Technology in exercise-based cancer rehabilitation: a cross-sectional study of receptiveness and readiness for e-Health utilization in Danish cancer rehabilitation. Acta Oncol 2019 May;58(5):610-618. [doi: 10.1080/0284186X.2018.1562213] [Medline: 30698060]
16. Thorsen IK, Rossen S, Glümer C, Midtgaard J, Ried-Larsen M, Kayser L. Health technology readiness profiles among Danish individuals with type 2 diabetes: cross-sectional study. J Med Internet Res 2020 Sep 15;22(9):e21195 [FREE Full text] [doi: 10.2196/21195] [Medline: 32930669]
17. Nielsen AS, Hanna L, Larsen BF, Appel CW, Osborne RH, Kayser L. Readiness, acceptance and use of digital patient reported outcome in an outpatient clinic. Health Informatics J 2022 Jun 03;28(2):14604582221106000 [FREE Full text] [doi: 10.1177/14604582221106000] [Medline: 35658693]
18. Terp R, Kayser L, Lindhardt T. Older patients' competence, preferences, and attitudes toward digital technology use: explorative study. JMIR Hum Factors 2021 May 14;8(2):e27005 [FREE Full text] [doi: 10.2196/27005] [Medline: 33988512]
19. Wahl AK, Osborne RH, Langeland E, Wentzel-Larsen T, Mengshoel AM, Ribu L, et al. Making robust decisions about the impact of health education programs: Psychometric evaluation of the Health Education Impact Questionnaire (heiQ) in diverse patient groups in Norway. Patient Educ Couns 2016 Oct;99(10):1733-1738. [doi: 10.1016/j.pec.2016.05.001] [Medline: 27211224]
20. Wahl AK, Hermansen A, Osborne RH, Larsen MH. A validation study of the Norwegian version of the Health Literacy Questionnaire: A robust nine-dimension factor model. Scand J Public Health 2021 Jun;49(4):471-478 [FREE Full text] [doi: 10.1177/1403494820926428] [Medline: 32508258]
21. Terp R, Lindhardt T, Kayser L. Theory-driven development of an educative nutritional intervention (ENI) supporting older hospital patients to eat sufficiently, assisted by an eHealth solution: an intervention mapping approach. BMC Health Serv Res 2022 Nov 28;22(1):1435 [FREE Full text] [doi: 10.1186/s12913-022-08679-8] [Medline: 36443804]
22. Krag T, Jørgensen E, Phanareth K, Kayser L. Experiences with in-person and virtual health care services for people with chronic obstructive pulmonary disease: qualitative study. JMIR Rehabil Assist Technol 2023 Aug 14;10:e43237 [FREE Full text] [doi: 10.2196/43237] [Medline: 37578832]
23. Larsen MH, Strumse YS, Andersen MH, Borge CR, Wahl AK. Associations between disease education, self-management support, and health literacy in psoriasis. J Dermatolog Treat 2021 Sep;32(6):603-609 [FREE Full text] [doi: 10.1080/09546634.2019.1688233] [Medline: 31692398]
24. Hermansen A, Andersen MH, Borge CR, Dahl KG, Larsen MH, Lønning K, et al. Preliminary validity testing of the eHealth Literacy Questionnaire (eHLQ): a Confirmatory Factor Analysis (CFA) in Norwegian hospitalized patients. BMC Psychol 2023 Nov 23;11(1):409 [FREE Full text] [doi: 10.1186/s40359-023-01449-z] [Medline: 37996858]
25. Kayser L, Karnoe A, Duminski E, Jakobsen S, Terp R, Dansholm S, et al. Health professionals' eHealth literacy and system experience before and 3 months after the implementation of an electronic health record system: longitudinal study. JMIR Hum Factors 2022 Apr 29;9(2):e29780 [FREE Full text] [doi: 10.2196/29780] [Medline: 35486414]
26. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
27. Ward JH. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 1963 Mar;58(301):236-244. [doi: 10.1080/01621459.1963.10500845]
28. Forgy E. Cluster analysis of multivariate data fficiency versus interpretability of classifications. Biometrics 1965;21(3):768-769.
29. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572-572 [FREE Full text] [doi: 10.1136/bmj.314.7080.572] [Medline: 9055718]
30. De Moortel D, Palència L, Artazcoz L, Borrell C, Vanroelen C. Neo-Marxian social class inequalities in the mental well-being of employed men and women: the role of European welfare regimes. Soc Sci Med 2015 Mar;128:188-200. [doi: 10.1016/j.socscimed.2015.01.027] [Medline: 25621400]
31. Dreger S, Buck C, Bolte G. Material, psychosocial and sociodemographic determinants are associated with positive mental health in Europe: a cross-sectional study. BMJ Open 2014 May 28;4(5):e005095 [FREE Full text] [doi: 10.1136/bmjopen-2014-005095] [Medline: 24871540]
32. Lukaschek K, Vanajan A, Johar H, Weiland N, Ladwig K. "In the mood for ageing": determinants of subjective well-being in older men and women of the population-based KORA-Age study. BMC Geriatr 2017 Jun 16;17(1):126 [FREE Full text] [doi: 10.1186/s12877-017-0513-5] [Medline: 28622764]
33. Andersen MH, Hermansen A, Dahl KG, Lønning K, Meyer KB, Vidnes TK, et al. Profiles of health literacy and digital health literacy in clusters of hospitalised patients: a single-centre, cross-sectional study. BMJ Open 2024 May 20;14(5):e077440 [FREE Full text] [doi: 10.1136/bmjopen-2023-077440] [Medline: 38772592]
34. Dalmaijer ES, Nord CL, Astle DE. Statistical power for cluster analysis. BMC Bioinformatics 2022 May 31;23(1):205 [FREE Full text] [doi: 10.1186/s12859-022-04675-1] [Medline: 35641905]

Abbreviations

eHLQ: eHealth Literacy Questionnaire

heiQ: Health Education Impact Questionnaire
HLQ: Health Literacy Questionnaire
READHY: Readiness and Enablement Index for Health Technology
T2DM: type 2 diabetes mellitus
WHO-5: World Health Organization-Five Well-Being Index

Edited by J Wang; submitted 05.06.24; peer-reviewed by S Kujala, I de Siqueira Silva, L Donelle, X Huang; comments to author 04.07.24; revised version received 26.07.24; accepted 14.01.25; published 07.02.25.

Please cite as:
Bergh S, Benth JŠ, Høgset LD, Rydjord B, Kayser L
Assessment of Technology Readiness in Norwegian Older Adults With Long-Term Health Conditions Receiving Home Care Services: Cross-Sectional Questionnaire Study
JMIR Aging 2025;8:e62936
URL: https://aging.jmir.org/2025/1/e62936
doi:10.2196/62936
PMID:39918862

Original Paper

Experiences of Older Mental Health Patients and Their Care Partners Using a Proxy Account to Access Open Notes: Qualitative Interview Study

Eva Meier-Diedrich[1,2], MSc; Tobias Esch[3], MD; Maria Hägglund[4,5], PhD; Martin Heinze[1,2,6], MD; Stefan Hochwarter[7], PhD; Justin Speck[2], BSc; Marie Wagener[2], MSc; Volker Dahling[1,2,8*], MD; Julian Schwarz[1,2,6*], MD

[1]Department of Psychiatry and Psychotherapy, Center for Mental Health, Immanuel Hospital Rüdersdorf, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[2]Faculty of Health Sciences Brandenburg, Brandenburg Medical School Theodor Fontane, Neuruppin, Germany

[3]Institute for Integrative Health Care and Health Promotion, Faculty of Health, School of Medicine, Witten/Herdecke University, Witten, Germany

[4]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

[5]Uppsala University Hospital, Uppsala, Sweden

[6]Center for Health Service Research Brandenburg, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[7]Institute for Biomedicine and Health Sciences (HEALTH), Joanneum Research Forschungsgesellschaft m.b.H, Graz, Austria

[8]Research Group Geriatric Psychiatry Research, Brandenburg Medical School Theodor Fontane, Rüdersdorf, Germany

[*]these authors contributed equally

Corresponding Author:
Eva Meier-Diedrich, MSc
Department of Psychiatry and Psychotherapy, Center for Mental Health
Immanuel Hospital Rüdersdorf
Brandenburg Medical School Theodor Fontane
Seebad 82/83
Rüdersdorf, 15562
Germany
Phone: 49 33638 83 501
Email: Eva.Meier-Diedrich@mhb-fontane.de

Abstract

Background: Older patients with serious mental illnesses such as cognitive disorders often rely on family members or spouses (care partners [CPs]) to meet their health care needs. CPs frequently lack essential information to fully understand the patients' illnesses and effectively support their treatment. Open Notes provide patients with digital access to their health care professionals' clinical notes and are associated with many positive outcomes, such as increased adherence and empowerment. However, older patients who use Open Notes may encounter use barriers such as limited digital literacy. Recent developments allow CPs to access Open Notes (proxy access) and receive valuable information, which holds significant potential for improving the care of older patients.

Objective: This study explored the experiences, barriers, and opportunities of older mental health patients and their CPs related to using Open Notes. Furthermore, influencing factors and interdependencies were identified.

Methods: Older patients (n=10) and their CPs (n=10) were provided with web-based proxy access to clinical documentation through a web-based patient portal. In-depth qualitative interviews (N=20) were conducted to explore experiences with this access. Data analysis was conducted in accordance with the constructivist grounded theory approach.

Results: The prerequisites for using Open Notes with proxy access were sufficient digital literacy on the part of the patient or CP, as well as the establishment of a trusting relationship between patients and CPs. Access to Open Notes enabled patients and CPs to gain a deeper understanding of the illness and its treatment while also facilitating enhanced contact with health care professionals. This resulted in greater involvement in the treatment process but may also prompt changes in relationship dynamics—CPs are better equipped to support patients in their health care but may also tend to monitor or control them through Open Notes. As a result, the introduction of Open Notes was accompanied by mixed feelings.

Conclusions: It is of utmost importance to provide older patients with comprehensive access to Open Notes to preserve their health autonomy. However, the involvement of CPs through proxy access is of great value in improving the care of older patients, especially those with cognitive impairments.

(JMIR Aging 2025;8:e66690) doi:10.2196/66690

KEYWORDS

psychiatry; eHealth; mental health; digital literacy; older patients; older adult; care partner; proxy access; open record access; Open Notes; patient portal; artificial intelligence; AI

Introduction

Background

The ongoing digitalization throughout society, coupled with demographic transformations, presents both opportunities and challenges for the health care sector. The global population is living longer, and in particular in industrialized countries, fertility rates are declining, which is significantly altering the age composition of the population [1]. Concurrently, digitalization is gradually permeating all areas of society, including the mental health care sector, thereby providing patients with new opportunities to participate in their treatments [2]. The European Health Data Space proposal follows this development and aims to improve patients' access and control over their personal electronic health data [3]. One such opportunity for patient participation is through online record access (ORA), which enables patients to view their electronic health records (EHRs) via web-based patient portals. When patients are able to also access their health care professionals' (HCPs) notes via ORA, this is referred to as Open Notes [4-6]. In the United States and some Scandinavian countries, Open Notes are already an established practice in health care [3,7]. Patients in the United Kingdom have the option of accessing the clinical notes of their general practitioners [8]. In contrast, in other countries such as Germany, Open Notes have yet to be integrated into the health care system [9,10]. However, current legislation requires all hospitals in Germany to implement patient portals. Therefore, in the near future, it will be technically possible to offer Open Notes [11]. Access to clinical notes and the associated transparency can enhance patient empowerment and engagement, improve medication adherence, and bolster disease management and awareness [12-15]. Studies indicate that vulnerable groups, including older patients, particularly benefit from Open Notes [16]. Furthermore, it can enhance communication between (older) patients and their HCPs, thereby strengthening trust in the treatment process [17,18]. While many older patients express interest in and intent to use digital health technologies, few actually use them [19]. Because older patients are not considered digital natives, they often face several challenges when using digital health technologies, including limited digital health literacy, usability issues, and heightened concerns about data security [19-21]. In addition, older patients are increasingly affected by cognitive deficits (eg, dementia), which further complicates the development of digital health literacy [22]. These barriers lead to a decrease in the use of web-based health services such as ORA with advancing age, especially without (human) guidance [23]. Therefore, it is imperative that digital health literacy is taken into account in the design of patient portals and EHRs to ensure their accessibility and inclusivity for older users [24,25].

In general, a significant proportion of older patients rely on relatives such as family members, partners, or friends (referred to as care partners [CPs]) to meet their health care needs [24]. This is also true for digital health services [26]. Receiving support from CPs has a positive impact on patients' quality of life, quality of care, and health resource use. In addition, patients and CPs would like to have greater involvement in medical care, but this has not been adequately supported by the health care system [24,27]. CPs often lack essential information about the health status and treatment planning of older patients, which can significantly complicate care [24]. With proxy access, Open Notes allow relatives to read medical treatment documentation with the patient's consent [28]. Giving CPs access to Open Notes provides them with important information and facilitates their care. In addition, studies show that the involvement of CPs can also increase engagement in the treatment of (older) patients using ORA [29]. Current research confirms that such access is desired by CPs [22,30] but acceptance is still limited [26]. ORA and Open Notes will become increasingly important in health care, and the associated opportunities should be available to all patients [22,24]. Therefore, it is essential to gain a deeper understanding of the use patterns of older patients, the role of CPs, and the barriers and opportunities associated with them.

Objectives

To date, only a limited number of studies have examined the use of ORA by older patients in the context of proxy access [24,31,32]. Specifically, in the area of mental health, the authors are currently unaware of any studies on this topic [33,34]. In light of the current state of research, the purpose of this study was to explore the experiences, preferences, and needs of older mental health patients and their CPs, as well as the barriers and opportunities related to using Open Notes. In addition, this study aimed to provide recommendations for best practice in this area and sought to identify the factors influencing the impact of CP access to Open Notes.

Methods

Study Design

The *Piloting and evaluation of a participatory patient-accessible electronic health record for geriatric psychiatric patients and their care partners* (PEP.AGE) study is part of the *Piloting and evaluation of a participatory patient-accessible electronic health record in Psychiatry and Somatics* (PEPPPSY) project (2021-2026) [35,36]. In the PEPPPSY project, patients are provided with access to their HCPs' treatment and progress

notes via a dedicated patient portal [37]. Furthermore, the development and implementation of the patient portal are being examined from the dual perspectives of both patients and HCPs. The PEP.AGE study broadens the scope of the PEPPPSY target population by including not only the perspective of older patients but also that of their CPs. Given the exploratory nature of this study, a qualitative design was chosen to ensure a comprehensive and thorough examination of the use of Open Notes by older patients and their CPs.

Ethical Considerations

This study was approved by the ethics committee of Brandenburg Medical School Theodor Fontane (E-01-20210727) and registered with the German Clinical Trials Register (DRKS00030188). Participants were informed of the study content and procedures both verbally and in writing. Informed written consent was then obtained from all participants. All participants had the right to withdraw from the study at any time without any adverse consequences. All data were anonymized. Participants received a compensation of €40 (US $41.39) for their participation.

PEPPPSY App

The patient portal pilot, called PEPPPSY, was initiated as part of a research collaboration between the Norwegian University of Science and Technology and Brandenburg Medical School Theodor Fontane. The portal was developed through an ongoing iterative and participatory process [35,36]. In addition to accessing clinical notes, patients and CPs can respond to HCPs' entries with comments. HCPs are then notified of these comments and can respond to patients (or CPs) within the same thread. In the current second phase of the project, the pilot has been expanded to include access for CPs (proxy access), which will increase the accessibility and utility of the portal for a broader patient population.

Study Setting

This study was conducted in 2 psychiatric outpatient clinics (Rüdersdorf and Strausberg) of the Immanuel Hospital Rüdersdorf in the state of Brandenburg, Germany. Psychiatric outpatient clinics are specialized facilities that provide psychiatric care to patients with severe mental illness. These patients typically require comprehensive and multidisciplinary psychiatric treatment and often lack access to adequate care in other outpatient settings (such as psychiatric or general medical practices) due to the severity or chronicity of their psychiatric conditions.

Recruitment

Eligible participants were enrolled in the PEP.AGE study from June 2023 to January 2024. Participating patients had to be aged ≥60 years, receive treatment at 1 of the 2 designated sites, and be able to provide informed consent. Patients with risk factors such as self-harm or harm to others and severe cognitive impairment were excluded from the study. Participating CPs had to be adults (aged ≥18 years) and able to provide informed consent.

Data Collection

At enrollment, sociodemographic data were collected from both patients and their CPs. Patients and CPs were then introduced to the use of the patient portal by their HCP or a member of the study team. They were provided with a comprehensive, user-friendly manual and the option of one-on-one assistance to set up their accounts and learn how to use Open Notes step by step. Subsequently, the older patients and their CPs (with patient consent) were given access to the patient portal and the HCPs' clinical notes. At the beginning of the intervention phase, participants were randomly contacted to identify and address any barriers to use. At the end of the 3-month intervention phase, semistructured interviews were conducted with patients and CPs using previously developed interview guides to gain deeper insights into their actual experiences with Open Notes (Multimedia Appendix 1). The interviews lasted between 20 and 30 minutes each. Throughout the study (onboarding phase, intervention phase, and interview phase), the study team kept field notes documenting observations and contextual information [38].

Data Analysis

The qualitative interviews were audio recorded, pseudonymized, transcribed, and analyzed by 2 researchers with the computer-assisted analysis software MAXQDA (VERBI GmbH) using the constructivist grounded theory by Charmaz [39]. The selected analytical approach was appropriate to the research subject as this exploratory study aimed to iteratively develop theoretical concepts from the data. In accordance with the approach by Charmaz [39], the data analysis was conducted continuously, commencing with the earliest data gathering (initial interview). The interviews were initially coded line by line to facilitate the conceptualization of ideas and the development of preliminary codes. Subsequently, focused coding was conducted, whereby the most significant and frequent codes were identified, sorted, and synthesized into overarching categories. Following this, relationships between the categories were identified and connected into coherent theoretical concepts (theoretical coding). On the basis of the developed concepts and emerging theory, the research team returned to the field and gathered additional data on specific themes until theoretical saturation was achieved. For quality assurance purposes, the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist was used (Multimedia Appendix 2).

Results

Sociodemographic Data

A total of 10 patients and 10 CPs were interviewed via telephone (by EMD or MW), and their complete sociodemographic data are shown in Table 1. In total, 5 dyads (each consisting of a patient and their respective CP), as well as 5 independent patients and 5 independent CPs, were interviewed. All participants self-identified as White individuals, were born in Germany, and spoke German as their native language. The age of the patients ranged from 62 to 81 years, with a mean of 71.60 (SD 6.43) years. A total of 70% (7/10) of the patients identified as female, and 30% (3/10) identified as male. Most patients

(8/10, 80%) were retired, whereas a minority (2/10, 20%) were still employed. All CPs (10/10, 100%) were family members of the patients (mainly spouses or children). The ages of the CPs were more diverse, ranging from 45 to 81 (mean 61.20, SD 11.02) years. This was due to the participation of spouses (5/10, 50%), children (4/10, 40%), and other family members (1/10, 10%) of the patients as CPs. In total, 40% (4/10) of the CPs were already retired, whereas 60% (6/10) were still employed part time or full time. Before the start of the PEP.AGE study, the vast majority (8/10, 80%) of the older patients had already given their relatives access to their health information. Patients reported medical discussions with their HCPs to their CPs, shared medical correspondence and medication schedules with them, or were accompanied by CPs to medical appointments.

Table 1. Characteristics of the patient–care partner dyads.

Dyad number	Patients			Care partners		
	Age (y)	Sex	Diagnosis (*ICD-10*[a] code)	Age (y)	Relationship to patient	Employment status
1	67	Female	Mild cognitive disorder (F06.7)	62	Spouse	Part time
2	75	Female	Dementia in Alzheimer disease with late onset (F00.1)	51	Child	Full time
3	80	Male	Dementia in Alzheimer disease with late onset (F00.1)	69	Spouse	Retired
4	81	Female	Severe depressive episode without psychotic symptoms (F32.2)	45	Child	Full time
5	67	Female	Social phobias (F40.1)	69	Spouse	Retired
6[b]	69	Female	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	52	Child	Retired
7[b]	62	Male	Generalized anxiety disorder (F41.1)	53	Child	Full time
8[b]	68	Male	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	81	Spouse	Retired
9[b]	78	Female	Bipolar affective disorder, current episode mild or moderate depression (F31.3)	69	Spouse	Retired
10[b]	69	Female	Recurrent depressive disorder, current episode severe without psychotic symptoms (F33.2)	61	Other family member	Full time

[a]ICD-10: International Classification of Diseases, 10th Revision.

[b]No dyad; independent care partners and patients.

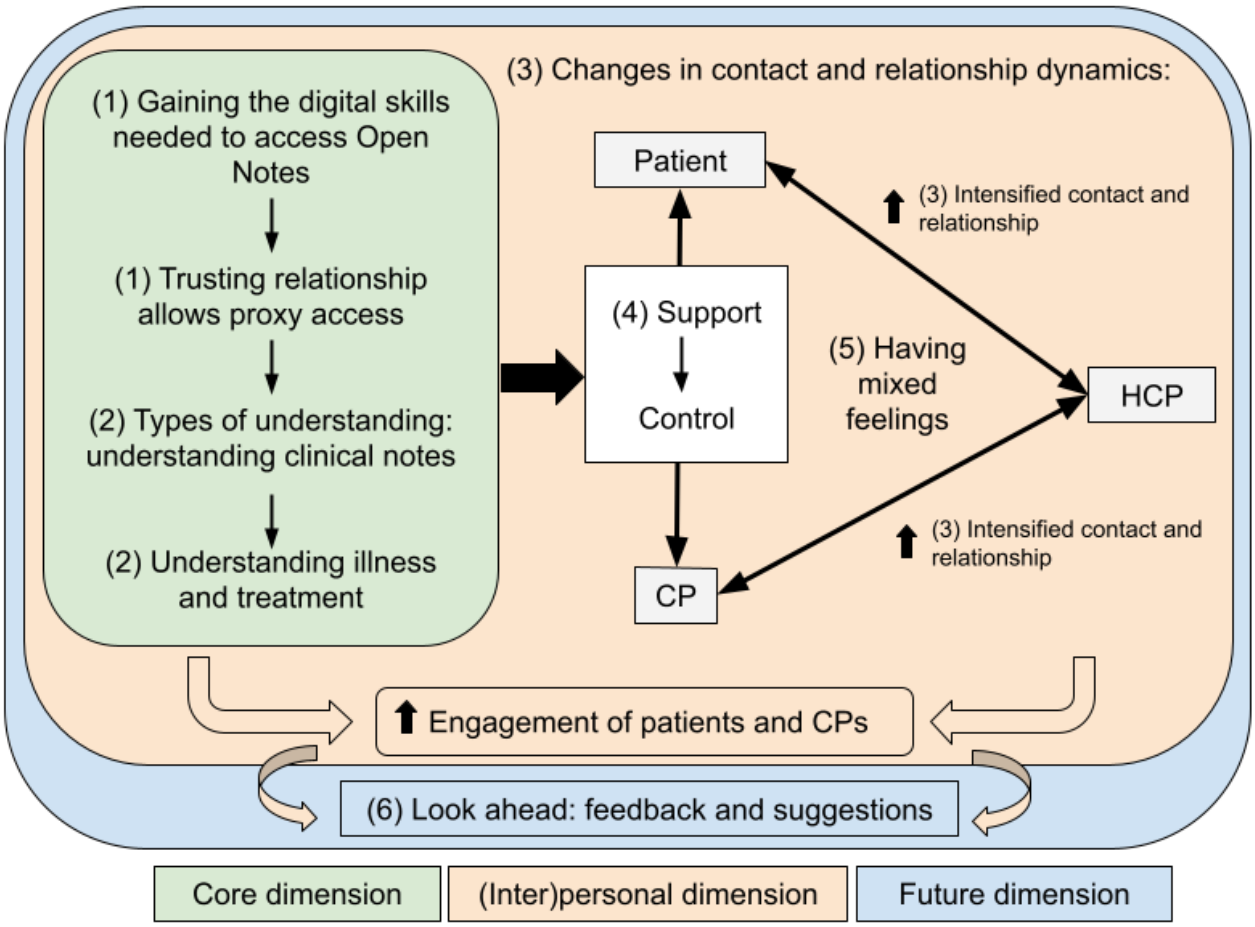
The patients supported by the independent CPs (CPs 6-10) were aged between 71 and 86 (mean 80.8, SD 6.099) years and retired, and they self-identified as White individuals. Patients were being treated for the following main diagnoses: dementia in Alzheimer disease with late onset (*International Classification of Diseases, 10th Revision* [*ICD-10*], code F00.1); mild cognitive disorder (*ICD-10* code F06.7); recurrent depressive disorder, current episode severe without psychotic symptoms (*ICD-10* code F33.2); and mixed anxiety and depressive disorder (*ICD-10* code F41.2).

Qualitative Findings

Overview

The results revealed 3 (partially interrelated) dimensions associated with Open Notes with proxy access when used by older patients and their CPs. These dimensions and their interactions are summarized in Figure 1 and described in detail in the following sections. The green core dimension in Figure 1 provides the foundation for the use of Open Notes and proxy access (eg, digital skills and literacy, trust, and understanding of note content). The red (inter)personal dimension encompasses the impact that Open Notes can have on the relationships among patients, CPs, and HCPs. Finally, the blue future dimension offers ideas and recommendations for the further development of Open Notes. Quotations from CPs and patients are identified using the IDs *CP* and *PAT*, respectively.

Figure 1. Interrelations of the qualitative categories and dimensions related to the use of Open Notes with proxy access by older patients and their care partners (CPs). Arrows pointing upward indicate increase or enhancement. HCP: health care professional.



Gaining the Digital Skills Needed to Access Open Notes

The interviews and field observations revealed that the digital health literacy of the older patients and their CPs varied widely—both between dyads (patient-care-partner pairs) and across participant groups. While some dyads required direct assistance from study team members to activate, log into, and use the patient portal, others required no assistance at all. Younger CPs generally found it much easier to navigate the patient portal than older patients. This variability was also evident in the interviews, with approximately half (4/10, 40%) of the patients and most of the CPs (6/10, 60%) reporting that using Open Notes was challenging or even beyond the digital literacy of the patients, as illustrated by the following statements:

> *It [technical difficulties] could very well be because I'm no longer able to do things like that. [I had] Two strokes and then the nerve disease.* [PAT_0420]

> *New technologies are often a challenge for older people. It takes a lot of patience and support from us younger ones, but it's doable if you stick with it and show them they can do it.* [CP_0213]

As the previous quotes illustrate, a significant number of patients (5/10, 50%) relied on the assistance of their CPs to use the patient portal and Open Notes. At the same time, 20% (2/10) of the CPs themselves indicated that using digital health technologies posed a (manageable) challenge to their digital literacy:

> *I had to overcome a few technical hurdles, but with time and some support, I managed to use the application.* [CP_0213]

The 2-factor authentication log-in process proved to be particularly challenging, if not insurmountable, for some patients and CPs. Some participants were able to overcome this barrier on their own or with assistance (from CPs or the study team). However, others became so frustrated that they stopped using the patient portal altogether. CPs reported the following:

> *Well, there's this two-step login process, where you need the SMS PIN, and then—since I mostly used it on mobile—you have to fiddle around a bit to find where the access to the records is, but ultimately it was okay.* [CP_0209]

> *I had a question once because the access didn't work at the beginning, but it was resolved relatively quickly. They sent me a new one, and I was able to use it. Okay, but I don't know exactly if it was my fault or if it was just issues with the program.* [CP_0409]

Although most participants reported using and perceiving benefits from the patient portal, 30% (3/10) of the CPs and 10% (1/10) of the patients indicated that they did not access the patient portal in their daily lives. This was due to a lack of

perceived need to review the information given their regular contact with the HCP and the absence of a crisis situation that would have made accessing the records more relevant:

> *For us, it [Open Notes] doesn't have any everyday use. The idea that there's the possibility to look up and comment on disputed questions is great. But in the six months we've been participating, there hasn't been a situation that required us to intervene or do anything.* [CP_0211]

Most patients (8/10, 80%) agreed that a solid foundation of trust is a prerequisite for disclosing sensitive clinical information to family members. Some CPs (3/10, 30%) expressed a similar view. An open and honest conversation about the advantages and disadvantages regarding the patients' privacy seemed to be particularly important before using Open Notes with a proxy access, as the following quote shows:

> *We had discussed beforehand what this is and what it means. Otherwise, I don't think she [CP] would have agreed to it. When everything is disclosed, you have to be willing to accept that. Some might say, "Oh no, I don't want that, it's too private," depending on who it is and the relationship involved.* [PAT_0419]

Types of Understanding

Overview

The implementation of Open Notes involves several types of understanding. First, it is essential for patients and CPs to comprehend the content of the clinical note, which requires that the notes be written in patient-friendly language. In addition, by reading the clinical notes, older patients and their CPs were able to gain a better understanding of the illness and its treatment. This allowed them to adequately prepare for medical appointments, reducing anxiety and facilitating understanding during treatment sessions.

Understanding Clinical Notes

Most interviewed patients (8/10, 80%) and half (5/10, 50%) of the CPs agreed that the documentation was particularly understandable when it was composed in a manner accessible to patients without a lot of medical or technical jargon:

> *It [Open Notes] was understandable. Without any medical jargon, everything was fine. The way I described it, he [the HCP] wrote it down, more or less in my own words. It was expressed a bit better, but still in normal, understandable terms, I would say.* [PAT_0411]

> *Yes, it was understandable. Of course, there are always medical terms that might be unfamiliar to a layperson at first. But then you remember the conversation and can figure out what it was about and what was meant. So far, I can't say that I didn't understand anything. It was all very understandable.* [CP_0201]

The previous quotes show that, for some users, the information in the Open Notes alone is not sufficient; rather, understanding is built by combining the knowledge gained from the conversation with HCPs during in-person medical appointments and the information provided in the Open Notes.

Some CPs accompanied patients to their medical appointments and were able to recall the content of the conversations. However, for CPs who were unable to attend appointments, the clinical notes seemed to be easy to understand, as the following quote illustrates:

> *Yes, those were his notes. Brief and to the point. Of course, he didn't write long texts, but at least he documented briefly what was discussed, how the medications are, and what the plan is going forward. He wrote it down in a way that was understandable for everyone.* [CP_0416]

Understanding the Illness and Its Treatment

Some patients (3/10, 30%), but especially CPs (6/10, 60%), reported that access to patient-friendly clinical notes provided them with a more complete understanding and awareness of the patient's illness and treatment:

> *I now have a much better overview of the entire treatment process and my mother's current health condition. This makes it easier to make informed decisions and plan the next steps.* [CP_210]

This improved understanding of the illness appeared to serve as a foundation for subsequent developments, including the increased involvement of CPs in the treatment process and the provision of support. Because treatment appointments often leave little time for questions or repetition, patients (5/10, 50%) and CPs (3/10, 30%) found open-ended notes to be a valuable reminder, allowing them to prepare for and follow up on appointments more easily. Participants found it beneficial to have a written record of what was discussed that they could review at their own pace, allowing them to process the information in a way that best met their individual needs:

> *Remembering and understanding important details and conversation points better. This was especially useful for preparing for appointments and following up on recommendations.* [CP_0213]

> *Because for me, it's easier when I see something in writing, read it, and then respond or share my own experiences. This back-and-forth, this exchange with the doctors and staff, it's easier for me in writing than sitting in front of the doctor who might not have much time.* [PAT_0402]

> *If we have an appointment, we review the last one together, summarize the key points, and build the new medical appointment on that.* [CP_0201]

As the preceding quotes illustrate, one patient noted that reviewing the documentation allowed her to better assess and understand her own treatment progress. Another patient reported that he used the clinical notes to confirm that his HCP had understood him correctly during the visit. This gave him the opportunity to address any potential misunderstandings:

> *Exactly, it allows you to see for yourself that everything was conveyed clearly. When you only attend the doctor's visit and then leave, you forget*

> *half of it anyway. It was much better for me to be able to read it again and confirm that I was truly understood.* [PAT_0204]

Changes in Contact and Relationship Dynamics

Nevertheless, many patients (7/10, 70%) and CPs (5/10, 50%) described Open Notes as facilitating communication with HCPs in a variety of ways. Open Notes, with the opportunity to comment on notes and send messages back and forth, made contact with HCPs faster, more direct, more efficient, more accessible, and more frequent. As a result, participants reported an increased sense of involvement in the treatment process and improved collaboration with HCPs:

> *You can communicate quickly. That is important. It's nice, it's accessible.* [CP_0409]

> *Yes, the contact became more frequent, and I felt more involved in the treatment. It was a very positive effect that improved collaboration.* [CP_0210]

In addition, Open Notes allowed participants to contact HCPs outside of office hours, which was particularly convenient for full-time CPs. All participants were aware that this was asynchronous communication and that responses would only be made during the HCPs' working hours. One patient said the following:

> *You can exchange messages, even on the quick...So having a direct line to the doctor [through Open Notes], without having to call during office hours, is relatively quick. You get a prompt response from the doctor to what you write.* [PAT_0204]

In some cases, communication via Open Notes even replaced telephone contact. This was seen as a relief by some participants (1/10, 10% of patients and 4/10, 40% of CPs) as HCPs were often difficult to reach by phone. In this context, one CP highlighted the portal's communication and commenting function as a valuable tool for patients with mental health conditions who may find phone calls challenging:

> *Interviewer: Okay. And compared to calling, did you feel that you could reach your healthcare provider better or faster through the patient portal than by phone or other means?...Patient: Definitely much better, because calling is always tricky. If they're in treatment, you can't reach them. But this way, they responded when they had time, and everything was handled very quickly, so it was totally fine.* [PAT_0204]

> *As I would say, when patients use it themselves, and I am not a patient, but patients generally have underlying issues, often psychological, which make it difficult for them to communicate. Yes, reaching for the phone is challenging, going to the doctor is difficult. But maybe writing is somewhat less personal and might be easier. And it can be done at night or at an inconvenient time without feeling guilty, so I can imagine this is definitely a good option that could continue to be used.* [CP_0409]

Moreover, a slight shift in the relationship dynamics between patients and their CPs was noted with the implementation of Open Notes with proxy access. Now that CPs had access to the clinical notes, some patients (4/10, 40%) felt that their CPs were more understanding of their mental illness. This suggests a developmental process on the part of the CPs initiated by reading the shared notes:

> *Well, my stepson initially had problems because he couldn't imagine it when I said that it's still hard for me to take the bus alone. I get such a racing heart...But as I said, my stepson couldn't really understand it. Maybe you can't fully understand if you are healthy. But he has learned to understand it.* [PAT_0224]

From Support to Control

There was considerable variability in the level of digital health literacy among the older patients (see the *Gaining the Digital Skills Needed to Access Open Notes* section). In particular, patients in the oldest age group and those with cognitive impairments relied heavily on their CPs to help them navigate the digital patient portal. Some patients delegated responsibility for managing their health information in the portal to their CPs alone, whereas others sought to collaborate with their CPs in reading and understanding clinical notes:

> *My mother is now 82, and at that age, she's not likely to engage with apps or registration issues. If anything, I managed it for her or we discussed it.* [CP_0419]

> *I can open this page and then we can read it together. Or I explain to him what my concerns were. But independently, no longer.* [CP_0201]

In contrast, a younger and more digitally literate patient reported that she was able to access her Open Notes independently and only sought assistance from her CP when she encountered issues:

> *Well, I would first read it [Open Notes] on my own because I only have my daughter as a relative. And she has her own problems at the moment, so I don't really need her help unless I have issues. I'll handle it myself first.* [PAT_0402]

The ability to access clinical notes enabled many CPs (6/10, 60%) to gain a deeper understanding of the illness and treatment of the family member with a severe mental health issue. This increased their confidence in providing effective care and managing the illness, allowing them to better support the treatment (such as preparing for medical appointments and adhering to medication plans):

> *It [Open Notes] significantly improved my understanding. I could better follow the treatment processes and medical decisions. This helped me support my mother better and make informed decisions.* [CP_210]

> *It [Open Notes] was sometimes difficult, especially when the reports were not positive. But it helped me be better prepared and respond quickly if something was wrong.* [CP_0213]

As previously indicated (see the *Changes in Contact and Relationship Dynamics* section), this heightened level of involvement and responsibility, in addition to the improved information flow, was partially attributable to more intensive contact with HCPs, as evidenced by the following quote:

> *Yes, the contact became more intense, and I felt more included in the treatment. It was a very positive effect that improved the collaboration.* [CP_0210]

A few CPs (3/10, 30%) reported that Open Notes enabled them to provide more effective support from a greater distance (eg, from another city or country). Furthermore, CPs observed that reading the clinical notes reduced the need to accompany patients to medical appointments, thereby enhancing autonomy for both CPs and patients:

> *This allowed me to monitor their health data and ensure they received the right care even when I couldn't be with her. It gave me a sense of security to always be informed.* [CP_0210]

In addition to increased involvement, responsibility, and support, some CPs (3/10, 30%) also demonstrated a tendency to monitor or control patients through Open Notes. They compared the patient-reported information from medical appointments with the written information to assess the veracity of the reports and identify any potential omissions. This monitoring held the potential for conflict, but in one case, it also led to a more open and honest exchange between a patient and a CP regarding their inner motivations (eg, the withholding of information due to feelings of embarrassment, fear of disempowerment, or memory issues) and, thereby, enhanced mutual understanding:

> *It [Open Notes] allowed me to access all relevant information and better monitor my mother's health. It helped us be better informed and respond more quickly to changes.* [CP_0210]
>
> *There was a moment of surprise when she didn't mention something or had forgotten, but it was actually helpful because it led to a discussion where I could address it. She was honest, and we could discuss things in more detail or I could suggest she pay more attention to certain aspects. So, it wasn't a bad thing; it facilitated further discussion.* [CP_0416]

Having Mixed Feelings

Both CPs (7/10, 70%) and patients (5/10, 50%) provided insights into their emotional perceptions of the Open Notes. Notably, both patients and CPs reported a similar range of emotional experiences, including both positive and negative feelings. Both groups reported feelings of emotional distress associated with reading about deteriorating health or lack of treatment success. One patient even described experiencing persistent worrying thoughts. Some patients (4/10, 40%) also expressed concern that their CPs might experience distress as a result of reading the notes:

> *Well, it's a bit burdensome, I would say, maybe. When something new comes up and then the success doesn't happen.* [PAT_0402]
>
> *It was a mix of relief and concern. Relief because I was informed, and concern when the information wasn't positive. But overall, it helped me to be better prepared.* [CP_0210]
>
> *I am concerned that my relative may be emotionally distressed by reading the entries [Open Notes].* [PAT_0207]

Both CPs and patients noted that Open Notes provided a sense of security regarding the illness and its treatment. This sense of security was derived from 2 sources: first, the ability to access treatment information and, second, the knowledge that this information has been validated by experts:

> *No, for me it's more like the lack of knowledge is stressful. When you have an informed status, you can handle it better.* [CP_0211]
>
> *Yes, actually good, because I know it comes from a competent source and not just from random internet readings where every third person says something different, and so on. So, for me, it's reliable information. Definitely, knowing without having to worry about whether it's true or not or maybe or something, so that's more reassuring for me.* [PAT_0402]

One patient found it motivating and encouraging to read her HCP's notes. In addition, this patient was particularly motivated by proxy access and the fact that her CP also read the notes, which led her to engage in more self-care. This particular finding suggests that access to clinical notes by CPs may also impact treatment outcomes or patient recovery on a personal level beyond the increased involvement of well-informed CPs in the patients' health management:

> *But when you read the family's comments, like, "Hey, you've been letting yourself go lately," or "You seem unmotivated," it motivates you. You realize they are right; there's no reason to just hang around or whatever.* [PAT_0411]

In contrast, for other patients (2/10, 20%) and CPs (2/10, 20%), the clinical notes were less emotionally significant and were perceived more neutrally, as illustrated by the following quote:

> *You perceive it relatively neutrally. You don't get super happy or deeply depressed.* [PAT_0204]

Looking Ahead: Suggestions and Feedback

Several patients and CPs provided feedback on potential modifications to the patient portal that could improve its usability. Typically, suggestions focused on modifying or enhancing existing features within the portal. For instance, 4 participants (n=3, 30% patients and n=1, 10% CPs) expressed a desire for a read receipt feature to confirm when HCPs had received and read their messages. In addition, 20% (2/10) of the patients proposed that they be notified when a response from their HCP had been submitted. This notification feature had already been incorporated into the system and could be enabled by the user, yet these patients were unaware of its availability:

> *Something like that, just like with emails where you can send a confirmation of receipt or read receipt,*

> *so you know it's been received and opened. Sometimes, that's all you need.* [CP_0416]
>
> *And if the HCP has written something, it would be nice to get an email notification so that I know there's a message there. If it's out of the ordinary and you don't check it every day, you might not see it for a few days.* [PAT_0204]

Moreover, one patient expressed a desire for the portal's features to be more appealing and engaging for older patients, with the goal of tailoring the portal's design to the needs of this patient group (eg, by encouraging them to write comments). Both improved guidance and an optimized design were requested by the participants and could facilitate greater accessibility and appeal for the target audience:

> *A little guidance, maybe. Okay, that we start here, with a specific topic being set. I need to know, what should I write?* [CP_0402]
>
> *But the commenting function should be designed in a way that makes you want to use it, that makes you feel like speaking up.* [PAT_0420]

Other participants expressed a preference for integrating additional features into the patient portal beyond simply reading clinical documentation. In total, 10% (1/10) of the patients and 20% (2/10) of the CPs suggested enriching the patient portal with more psychoeducational information about the illness and integrating some type of psychoeducational lexicon or psychiatric frequently asked questions into the patient portal:

> *And I would certainly wish for a way to learn more about the illness, about behavioral strategies, options for the CP, but also for the patient. So you don't have to Google and look for information elsewhere. If you're already in the psychiatric system, maybe you could listen to more. Do you understand? That on this platform, on this level, you could already have specific questions answered.* [CP_0201]

In fact, one CP expanded the original scope of Open Notes by using the commenting feature to document important developments in her mother's health. She used this primarily as a personal reminder while also indirectly facilitating transparency and understanding of progress for HCPs by posting it on the patient portal:

> *At the beginning, I would write down things that I noticed in my daily life with my mom, as a personal reminder. It was helpful to have these notes ready for the next appointment as preparatory points. I definitely find it useful for that.* [CP_0409]

Discussion

Synthesis of the Findings

The results highlight both opportunities and challenges associated with using Open Notes for older patients and their CPs. In addition, the key drivers of proxy access were identified, and their interdependencies were highlighted. Our results show that older patients and their (sometimes older) CPs must first gain (proxy) access to the patient portal to use and benefit from Open Notes. This requires sufficient digital literacy and mutual trust between patients and CPs. In our study, many older patients needed support from their CPs to navigate digital health services. Once access to open records was established, both parties reported feeling more informed about the illness and its treatment and more in touch with HCPs. These 2 factors led to increased health literacy, engagement, and involvement for both patients and CPs. In line with this, our results suggest that access to Open Notes enables CPs to better support patients in their (digital) health management. However, there was also evidence that CPs used Open Notes to control patients, which could lead to conflicts. Finally, recommendations for further developments and feedback emerged.

Our findings show that it is particularly valuable to allow patients, CPs, and HCPs to digitally engage with Open Notes via a comment function, allowing the stakeholders to directly communicate by sending asynchronous messages. This interactivity of the test environment (*PEPPPSY*) in which our study was conducted was frequently used and highly valued by patients and CPs. In particular, it appeared to contribute to the health literacy of patients and CPs by allowing them to ask questions about the content of the notes or the treatment in general. Communication via Open Notes is not a classic feature of Open Notes, nor is it simply a secure messaging function as the digital interaction is not separate from the Open Note itself. This demonstrates that Open Notes serve multiple purposes (such as providing information, facilitating contact, and offering reassurance) depending on the level of interactivity available [40].

Ensuring Accessibility for Older Patients

As evidenced by previous research and observed in our study as well, older patients (and their CPs) predominantly use Open Notes as a memory aid, benefiting from this tool to prepare for or recap medical appointments. This provides both patients and CPs with an increased sense of security in their treatment processes. The positive effects of the implementation of Open Notes with a proxy access shown in this study—such as enhanced patient empowerment and engagement, increased CP involvement, and improved health management—align with those found in previous research [30,41,42]. Nevertheless, for these advantages to be fully realized, the initial challenge must be addressed: ensuring that older patients and their occasionally also older CPs have convenient access to the patient portal. This seems particularly relevant as older adults show interest in using patient portals yet the existence of numerous barriers hinders their ability to do so [19]. While CPs can indeed play a crucial role in compensating for the patients' lack of digital health literacy—as observed in our study—it is equally important to encourage and enable older patients (with sufficient cognitive abilities) to independently access their health information. Older adults are often apprehensive or skeptical about digital health tools, so addressing these concerns is essential [43]. Furthermore, it is important to ensure that older patients are able to comprehend the content of the Open Notes. Consequently, the clinical documentation must be written in patient-friendly language, which is not always the case in clinical practice [6].

On the basis of the study results, it seems important to provide older patients and their CPs with a clear and detailed explanation of the available features (such as the commenting feature and the opt-in notification feature) before they use the patient portal and Open Notes. It is imperative that patient portal interfaces are designed in a manner that is accessible to all age groups and that the technical requirements are kept as user-friendly as possible. For instance, alternative methods of 2-factor authentication should be explored as requiring users to use 2 devices (eg, a phone and a computer) simultaneously can be overwhelming and frustrating and may, ultimately, result in older patients giving up on using the patient portal and Open Notes. At the same time, increased usability must be compatible with high-level data security requirements. Furthermore, the design of patient portal interfaces should adhere to fundamental age-specific design principles, including the use of appropriate fonts, color choices, and audio alternatives and the minimization of text entry requirements [44]. Providing users with the option to select either a *standard* or *older age–accessible* interface design when accessing the patient portal could prove advantageous as it would enable users to customize their experience to align with their specific requirements. Nevertheless, it seems unlikely that merely modifying the interface will be sufficient to significantly increase the adoption of patient portals and Open Notes among older patients, and therefore, a more comprehensive approach is needed.

First, older patients must be made aware of the availability of patient portals (and the possibility of setting up a proxy access) through comprehensive and targeted informational campaigns [45]. Second, older patients need to be encouraged to use patient portals through the aforementioned campaigns and, more importantly, through their general practitioners and other HCPs [46]. As highlighted in the interviews and supported by findings of other studies, human guidance is essential for older patients to use the full range of features available on patient portals [47]. This responsibility should not be borne solely by CPs, particularly given that not all older patients have access to a digitally literate CP [48].

In light of the ongoing digitalization of the health care system, it may be worthwhile to consider the introduction of an institutionalized role dedicated to this task. In the United States, the role of digital navigator is currently being investigated and defined [49]. Digital navigators are HCPs who have undergone specialized training in the area of digital mental health applications. They provide consultative assistance to health care providers and offer continuous guidance to patients in using these applications [50]. To date, this role has been primarily concerned with the use of digital health applications. However, in light of the increasing international adoption and promotion of EHRs and ORA, it may be beneficial to consider expanding the role of digital navigators to encompass these additional tools and consider the integration of artificial intelligence–assisted support—as proposed by Wunderlich et al [48] with the concept of digital case managers. Nevertheless, artificial intelligence solutions should not replace human guidance as one of the primary concerns of older patients is that digital services could potentially supplant personal contact with HCPs [43].

Preserving the Autonomy of Older Patients

In addition to the aforementioned fundamental requirements for adapting patient portals to ensure accessibility for older adults, other aspects must be considered when involving CPs through proxy access. In the course of our study, the themes of trust, support, and control emerged as particularly salient, a finding that is also corroborated by other studies [30]. Despite the fact that CPs are only granted proxy access with the informed consent of the older patients, there is a risk that they may use Open Notes as a tool for control. This issue has significant ethical and practical implications [51]. Patients with dementia are especially reliant on the assistance of CPs in the management of their health [52], which also applies to the use of digital patient portals [53,54]. It could be argued that, particularly in cases in which patients are experiencing significant cognitive decline, such as with dementia, a certain level of control may be necessary and appropriate within the context of their care. Moreover, all participating older patients consented to the involvement of their relatives (and may revoke this consent at any time), thereby indicating their general assent to the sharing of information (and, thus, also to the potential for control). Nevertheless, this controlling behavior represents a substantial limitation of the patients' autonomy and may potentially give rise to conflict in the relationship between patients and CPs. Furthermore, in accordance with the systemic *concept of a good reason*, it can be assumed that all behaviors, including the deliberate withholding of information by patients, are motivated by a good inner reason and represent a more or less constructive coping mechanism in the face of challenges and difficulties (eg, to avoid shame or to maintain independence) [55].

At this point, it is pertinent to re-examine whether these considerations are applicable to patients with age-related cognitive impairments. These reflections could likely be extended indefinitely, leading to a vicious circle. However, it is possible to diverge from this loop and conclude that it is essential to preserve the dignity of older patients (with and without dementia) while using ORA and Open Notes [56]. Therefore, it is crucial to consider how the experience of (controlling and) being controlled can develop into a trusting dependence on the support of CPs [57]. Caine et al [58] and Latulipe et al [30] suggest that patients should be informed precisely about which information CPs can access in the patient portal. It seems particularly important to provide patients with the option of fine-grained access settings, allowing them to decide which information should be shared and which should not [59]. Such fine-grained functionality was available in early versions of the Swedish national patient portal [3,60]. In addition, a *break-glass* access control protocol can be implemented whereby patients can define which information should be released in an emergency (eg, in the event of a significant deterioration in cognitive health or an unexpected hospitalization) [30,61]. Careful attention must be paid to defining the end points meticulously and distinctly (eg, establishing clear criteria for what constitutes significant cognitive decline).

Implications

Older patients can benefit significantly from Open Notes with proxy access in their health care. However, to realize these benefits, older patients (and their CPs) must first be empowered to access the patient portal and understand clinical documentation. This requires adapting the design of patient portals to the needs of the older patient population and supporting the digital literacy of older patients through tailored individual and structural interventions. Enabling patients and CPs to interact with their HCPs through Open Notes seems to be a particularly important new feature. Many older patients rely on the support of CPs to manage their health care, especially when using digital health services. However, to ensure that older patients maintain their autonomy and dignity when using digital health services such as Open Notes, it is crucial to prevent these tools from becoming instruments of control for CPs. Older patients should be able to make granular decisions about what information they want to share with their CPs and what they want to keep private. For emergencies, a *break-glass* access protocol should be established in advance.

Limitations

This study was based on a small number of participants, which limits the generalizability of the results. Furthermore, the group of participants was highly homogeneous with regard to the categories of race and migration status. To obtain generalizable results, larger studies with a more diverse selection of participants are required in the future. Similarly, older patients and their CPs represent a relatively specific participant group, which further limits the generalizability of the results. Younger patients and their CPs (eg, children or adolescents) may have differing user experiences and encounter completely different barriers and opportunities while using Open Notes. It is also necessary to consider the potential influence of social desirability bias. It can be assumed that older patients might have occasionally embellished their statements regarding their own digital literacy and the usability of the patient portal. For instance, a greater number of CPs than patients indicated that the patients experienced difficulties when using the portal. It is difficult to acknowledge one's own shortcomings and limitations. Furthermore, it is particularly challenging to do so in the presence of others, such as CPs and interviewers, as this could lead to embarrassment and perceived loss of status. It should also be noted that the interviews were conducted via telephone. Although the participants were asked to find a quiet and secure place, it cannot be guaranteed that they were undisturbed throughout the interview. However, the telephone interviews allowed the older patients to remain in their homes (ie, they did not have to travel long distances) and have quick access to support from their CPs in case of difficulties (eg, technical problems or comprehension problems due to cognitive deficits). Thus, telephone interviews allowed older patients with physical or cognitive impairments, as well as CPs living in other cities or countries, to take part in the study and reduced barriers to study participation.

Conclusions

Our study suggests that access to Open Notes can facilitate understanding and engagement between patients and their CPs and is associated with improved communication with HCPs. This may influence the dynamics of the triadic relationship among patients, CPs, and HCPs, with potential implications for power dynamics. In summary, no single patient portal can be expected to meet the needs of all patients—one size does not fit all. Individual solutions and adaptations of ORA are clearly needed to ensure acceptance and meaningful use by older patients and their HCPs.

Acknowledgments

This study was funded by the Brandenburg Medical School Theodor Fontane publication fund supported by the Ministry of Science, Research, and Culture of the State of Brandenburg.

Authors' Contributions

All authors contributed to planning, analysis, and the critical interpretation of the findings. JS and EMD contributed to the study design, and JS supervised the study. EMD and MW contributed to data collection. EMD conducted the data analysis. EMD wrote the first draft of the manuscript. EMD, VD, and JS modified successive drafts. JS edited and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guides.
[PDF File (Adobe PDF File), 125 KB - aging_v8i1e66690_app1.pdf]

Multimedia Appendix 2
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 15 KB - aging_v8i1e66690_app2.docx]

References

1. Bongaarts J. Human population growth and the demographic transition. Philos Trans R Soc Lond B Biol Sci 2009 Oct 27;364(1532):2985-2990 [FREE Full text] [doi: 10.1098/rstb.2009.0137] [Medline: 19770150]
2. Dal Mas F, Massaro M, Rippa P, Secundo G. The challenges of digital transformation in healthcare: an interdisciplinary literature review, framework, and future research agenda. Technovation 2023 May;123:102716. [doi: 10.1016/j.technovation.2023.102716]
3. Hägglund M, Kharko A, Bärkås A, Blease C, Cajander Å, DesRoches C, et al. A Nordic perspective on patient online record access and the European health data space. J Med Internet Res 2024 Jun 27;26:e49084 [FREE Full text] [doi: 10.2196/49084] [Medline: 38935430]
4. Schwarz J, Esch T. [OpenNotes in mental health? Pro]. Nervenarzt 2022 Nov;93(11):1163-1164 [FREE Full text] [doi: 10.1007/s00115-022-01357-4] [Medline: 35913587]
5. Esch T, Mejilla R, Anselmo M, Podtschaske B, Delbanco T, Walker J. Engaging patients through open notes: an evaluation using mixed methods. BMJ Open 2016 Jan 29;6(1):e010034 [FREE Full text] [doi: 10.1136/bmjopen-2015-010034] [Medline: 26826154]
6. Meier-Diedrich E, Lyckblad C, Davidge G, Hägglund M, Kharko A, McMillan B, et al. Impact of patient online record access on documentation: a scoping review. JMIR Preprints Preprint posted online on July 25, 2024 [FREE Full text] [doi: 10.2196/64762]
7. 21st century cures act: the office of the national coordinator: interoperability and information blocking final rule overview for psychiatrists. American Psychiatric Association. URL: https://www.psychiatry.org/psychiatrists/practice/practice-management/health-information-technology/interoperability-and-information-blocking [accessed 2023-10-16]
8. Blease C, Torous J, Dong Z, Davidge G, DesRoches C, Kharko A, et al. Patient online record access in English primary care: qualitative survey study of general practitioners' views. J Med Internet Res 2023 Feb 22;25:e43496 [FREE Full text] [doi: 10.2196/43496] [Medline: 36811939]
9. Schwarz J, Hoetger C, Pluschke LS, Muehlensiepen F, Schneider M, Thoma S, et al. Psychiatrists' perceptions of conditions and consequences associated with the implementation of open notes: qualitative investigation. BMC Psychiatry 2024 Jun 10;24(1):430 [FREE Full text] [doi: 10.1186/s12888-024-05845-6] [Medline: 38858711]
10. Meier-Diedrich E, Neumann K, Heinze M, Schwarz J. [Attitudes and expectations of psychological and medical psychotherapists towards open notes: analysis of qualitative survey responses]. Psychiatr Prax 2024 Oct;51(7):376-384. [doi: 10.1055/a-2320-8929] [Medline: 38810902]
11. Wurster F, Di Gion P, Goldberg N, Hautsch V, Hefter K, Herrmann C, et al. Roger's diffusion of innovations theory and the adoption of a patient portal's digital anamnesis collection tool: study protocol for the MAiBest project. Implement Sci Commun 2024 Jul 15;5(1):74 [FREE Full text] [doi: 10.1186/s43058-024-00614-8] [Medline: 39010236]
12. Hägglund M, McMillan B, Whittaker R, Blease C. Patient empowerment through online access to health records. BMJ 2022 Sep 29;378:e071531 [FREE Full text] [doi: 10.1136/bmj-2022-071531] [Medline: 36175012]
13. Blease CR, Walker J, Torous J, O'Neill S. Corrigendum: sharing clinical notes in psychotherapy: a new tool to strengthen patient autonomy. Front Psychiatry 2020;11:636411 [FREE Full text] [doi: 10.3389/fpsyt.2020.636411] [Medline: 33536954]
14. Bailey SC, Wallia A, Wright S, Wismer GA, Infanzon AC, Curtis LM, et al. Electronic health record-based strategy to promote medication adherence among patients with diabetes: longitudinal observational study. J Med Internet Res 2019 Oct 21;21(10):e13499 [FREE Full text] [doi: 10.2196/13499] [Medline: 31638592]
15. van Kessel R, Ranganathan S, Anderson M, McMillan B, Mossialos E. Exploring potential drivers of patient engagement with their health data through digital platforms: a scoping review. Int J Med Inform 2024 Sep;189:105513 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105513] [Medline: 38851132]
16. Blease C, Fernandez L, Bell SK, Delbanco T, DesRoches C. Empowering patients and reducing inequities: is there potential in sharing clinical notes? BMJ Qual Saf 2020 Oct;29(10):1-2. [doi: 10.1136/bmjqs-2019-010490] [Medline: 32188711]
17. Wang X, Zhao YC. Understanding older adults' intention to use patient-accessible electronic health records: based on the affordance lens. Front Public Health 2022;10:1075204 [FREE Full text] [doi: 10.3389/fpubh.2022.1075204] [Medline: 36761325]
18. Meier-Diedrich E, Wördemann J, Speck J, Turvey C, Weibezahl M, Schwarz J. How patients affected by a severe mental health condition and health care professionals communicate via a patient portal: qualitative analysis of secure messages. JMIR Preprints Preprint posted online on June 27, 2024.
19. Nahm ES, Zhu S, Bellantoni M, Keldsen L, Charters K, Russomanno V, et al. Patient portal use among older adults: what is really happening nationwide? J Appl Gerontol 2020 Apr;39(4):442-450 [FREE Full text] [doi: 10.1177/0733464818776125] [Medline: 29779422]
20. Hoogland AI, Mansfield J, Lafranchise EA, Bulls HW, Johnstone PA, Jim HS. eHealth literacy in older adults with cancer. J Geriatr Oncol 2020 Jul;11(6):1020-1022 [FREE Full text] [doi: 10.1016/j.jgo.2019.12.015] [Medline: 31917114]
21. Tinmaz H, Lee YT, Fanea-Ivanovici M, Baber H. A systematic review on digital literacy. Smart Learn Environ 2022 Jun 08;9:e13499. [doi: 10.1186/s40561-022-00204-y]
22. Antonio MG, Petrovskaya O, Lau F. Correction: the state of evidence in patient portals: umbrella review. J Med Internet Res 2021 Aug 16;23(8):e32421 [FREE Full text] [doi: 10.2196/32421] [Medline: 34398799]

23. Crouch E, Gordon NP. Prevalence and factors influencing use of internet and electronic health resources by middle-aged and older adults in a US health plan population: cross-sectional survey study. JMIR Aging 2019 Mar 26;2(1):e11451 [FREE Full text] [doi: 10.2196/11451] [Medline: 31518256]
24. Wolff JL, Dukhanin V, Burgdorf JG, DesRoches CM. Shared access to patient portals for older adults: implications for privacy and digital health equity. JMIR Aging 2022 May 04;5(2):e34628 [FREE Full text] [doi: 10.2196/34628] [Medline: 35507405]
25. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556 [FREE Full text] [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
26. Wolff JL, Berger A, Clarke D, Green JA, Stametz R, Yule C, et al. Patients, care partners, and shared access to the patient portal: online practices at an integrated health system. J Am Med Inform Assoc 2016 Nov;23(6):1150-1158. [doi: 10.1093/jamia/ocw025] [Medline: 27026614]
27. Dahling V, May S, Peter SV, Schwarz J, Ignatyev Y, Scheibner H, et al. [Expectations of patients and relatives for modern hospital care in geriatric psychiatry]. Z Gerontol Geriatr 2022 Oct;55(6):476-481. [doi: 10.1007/s00391-021-01941-4] [Medline: 34251463]
28. Wachenheim D. EHR shared access (Proxy) for family care partners. OpenNotes. URL: https://www.opennotes.org/proxy-access/ [accessed 2024-08-12]
29. Gleason KT, Peereboom D, Wec A, Wolff JL. Patient portals to support care partner engagement in adolescent and adult populations: a scoping review. JAMA Netw Open 2022 Dec 01;5(12):e2248696 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.48696] [Medline: 36576738]
30. Latulipe C, Quandt SA, Melius KA, Bertoni A, Miller DP, Smith D, et al. Insights into older adult patient concerns around the caregiver proxy portal use: qualitative interview study. J Med Internet Res 2018 Nov 02;20(11):e10524 [FREE Full text] [doi: 10.2196/10524] [Medline: 30389654]
31. Wolff JL, Darer JD, Larsen KL. Family caregivers and consumer health information technology. J Gen Intern Med 2016 Jan;31(1):117-121 [FREE Full text] [doi: 10.1007/s11606-015-3494-0] [Medline: 26311198]
32. Jackson SL, Mejilla R, Darer JD, Oster NV, Ralston JD, Leveille SG, et al. Patients who share transparent visit notes with others: characteristics, risks, and benefits. J Med Internet Res 2014 Nov 12;16(11):e247 [FREE Full text] [doi: 10.2196/jmir.3363] [Medline: 25405911]
33. Schwarz J, Meier-Diedrich E, Neumann K, Heinze M, Eisenmann Y, Thoma S. Reasons for acceptance or rejection of online record access among patients affected by a severe mental illness: mixed methods study. JMIR Ment Health 2024 Feb 05;11:e51126 [FREE Full text] [doi: 10.2196/51126] [Medline: 38315523]
34. Schwarz J, Bärkås A, Blease C, Collins L, Hägglund M, Markham S, et al. Sharing clinical notes and electronic health records with people affected by mental health conditions: scoping review. JMIR Ment Health 2021 Dec 14;8(12):e34170 [FREE Full text] [doi: 10.2196/34170] [Medline: 34904956]
35. Hochwarter S, Fehler G, Muente C, Eisenmann Y, Heinze M, Hibsch C, et al. Design of a patient-accessible electronic health record system in mental health. Stud Health Technol Inform 2022 May 25;294:583-584. [doi: 10.3233/SHTI220534] [Medline: 35612155]
36. Hochwarter S, Schwarz J. Facilitating collaboration by opening the common information space in mental health care. In: Proceedings of the 9th International Conference on Infrastructures for Healthcare. 2023 Presented at: InfraHealth 2023; September 11-12, 2023; Siegen, Germany. [doi: 10.1001/jamanetworkopen.2022.48696]
37. Piloting and evaluation of a participatory patient record in psychiatry and somatics (PEPPPSY). PEPPPSY. URL: https://www.pepppsy.de/ [accessed 2025-01-22]
38. Emerson RM, Fretz RI, Shaw LL. Writing Ethnographic Fieldnotes Second Edition. Chicago, IL: The University of Chicago Press; 2011.
39. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. Thousand Oaks, CA: SAGE Publications; 2006.
40. Berg M. Practices of reading and writing: the constitutive role of the patient record in medical work. Sociol Health Illn 2008 Jun 28;18(4):499-524 [FREE Full text] [doi: 10.1111/1467-9566.ep10939100]
41. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing telemedicine unreadiness among older adults in the United States during the COVID-19 pandemic. JAMA Intern Med 2020 Oct 01;180(10):1389-1391 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2671] [Medline: 32744593]
42. Reed ME, Huang J, Brand R, Ballard D, Yamin C, Hsu J, et al. Communicating through a patient portal to engage family care partners. JAMA Intern Med 2018 Jan 01;178(1):142-144 [FREE Full text] [doi: 10.1001/jamainternmed.2017.6325] [Medline: 29159402]
43. Landgren S, Cajander Å. Non-use of digital health consultations among Swedish elderly living in the countryside. Front Public Health 2021;9:588583 [FREE Full text] [doi: 10.3389/fpubh.2021.588583] [Medline: 34568247]
44. Liu N, Yin J, Tan SS, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501 [FREE Full text] [doi: 10.1093/jamia/ocab151] [Medline: 34472601]

45. Barysheva GA, Klemasheva EI, Nedospasova OP, Ngoc TT, Thang NC. Involvement of elderly people in the processes of modern digital transformations. Adv Gerontol 2022 Sep 11;12(3):324-330. [doi: 10.1134/S2079057022030043]
46. Latulipe C, Gatto A, Nguyen HT, Miller DP, Quandt SA, Bertoni AG, et al. Design considerations for patient portal adoption by low-income, older adults. Proc SIGCHI Conf Hum Factor Comput Syst 2015 Apr;2015:3859-3868 [FREE Full text] [doi: 10.1145/2702123.2702392] [Medline: 27077140]
47. Grossman LV, Masterson Creber RM, Benda NC, Wright D, Vawdrey DK, Ancker JS. Interventions to increase patient portal use in vulnerable populations: a systematic review. J Am Med Inform Assoc 2019 Aug 01;26(8-9):855-870 [FREE Full text] [doi: 10.1093/jamia/ocz023] [Medline: 30958532]
48. Wunderlich P, Wiegräbe F, Dörksen H. Digital case manager-a data-driven tool to support family caregivers with initial guidance. Int J Environ Res Public Health 2023 Jan 10;20(2):1215 [FREE Full text] [doi: 10.3390/ijerph20021215] [Medline: 36673969]
49. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
50. Perret S, Alon N, Carpenter-Song E, Myrick K, Thompson K, Li S, et al. Standardising the role of a digital navigator in behavioural health: a systematic review. Lancet Digit Health 2023 Dec;5(12):e925-e932 [FREE Full text] [doi: 10.1016/S2589-7500(23)00152-8] [Medline: 38000876]
51. Ganta T, Appel JM, Genes N. Patient portal access for caregivers of adult and geriatric patients: reframing the ethics of digital patient communication. J Med Ethics 2023 Mar;49(3):156-159. [doi: 10.1136/medethics-2021-107759] [Medline: 35437282]
52. Kasper JD, Freedman VA, Spillman BC, Wolff JL. The disproportionate impact of dementia on family and unpaid caregiving to older adults. Health Aff (Millwood) 2015 Oct;34(10):1642-1649 [FREE Full text] [doi: 10.1377/hlthaff.2015.0536] [Medline: 26438739]
53. Gleason KT, Wu MM, Wec A, Powell DS, Zhang T, Gamper MJ, et al. Use of the patient portal among older adults with diagnosed dementia and their care partners. Alzheimers Dement 2023 Dec;19(12):5663-5671 [FREE Full text] [doi: 10.1002/alz.13354] [Medline: 37354066]
54. Riffin C, Cassidy J, Smith JM, Begler E, Peereboom D, Lum HD, et al. Care partner perspectives on the use of a patient portal intervention to promote care partner identification in dementia care. J Appl Gerontol 2025 Jan;44(1):136-146. [doi: 10.1177/07334648241262649] [Medline: 38901834]
55. Kutz A. Systemische Haltung in Beratung und Coaching. Wiesbaden, Germany: Springer; 2020.
56. Tranvåg O, Petersen KA, Nåden D. Dignity-preserving dementia care: a metasynthesis. Nurs Ethics 2013 Dec;20(8):861-880. [doi: 10.1177/0969733013485110] [Medline: 23715554]
57. Wolfe SE, Greenhill B, Butchard S, Day J. The meaning of autonomy when living with dementia: a Q-method investigation. Dementia (London) 2021 Aug;20(6):1875-1890 [FREE Full text] [doi: 10.1177/1471301220973067] [Medline: 33372553]
58. Caine K, Kohn S, Lawrence C, Hanania R, Meslin EM, Tierney WM. Designing a patient-centered user interface for access decisions about EHR data: implications from patient interviews. J Gen Intern Med 2015 Jan;30 Suppl 1(Suppl 1):S7-16 [FREE Full text] [doi: 10.1007/s11606-014-3049-9] [Medline: 25480719]
59. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 01;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
60. Nurgalieva L, Cajander Å, Moll J, Åhlfeldt RM, Huvila I, Marchese M. 'I do not share it with others. No, it's for me, it's my care': on sharing of patient accessible electronic health records. Health Informatics J 2020 Dec;26(4):2554-2567 [FREE Full text] [doi: 10.1177/1460458220912559] [Medline: 32264734]
61. Marinovic S, Craven R, Ma J, Dulay N. Rumpole: a flexible break-glass access control model. In: Proceedings of the 16th ACM Symposium on Access Control Models and Technologies. 2011 Presented at: SACMAT '11; June 15-17, 2011; Innsbruck, Austria.

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
CP: care partner
EHR: electronic health record
HCP: health care professional
ICD-10: International Classification of Diseases, 10th Revision
ORA: online record access
PEP.AGE: Piloting and evaluation of a participatory patient-accessible electronic health record for geriatric psychiatric patients and their care partners
PEPPPSY: Piloting and evaluation of a participatory patient-accessible electronic health record in Psychiatry and Somatics

Edited by M Gray; submitted 20.09.24; peer-reviewed by F Wurster, EKO Wong; comments to author 15.10.24; revised version received 23.10.24; accepted 30.10.24; published 24.02.25.

Please cite as:
Meier-Diedrich E, Esch T, Hägglund M, Heinze M, Hochwarter S, Speck J, Wagener M, Dahling V, Schwarz J
Experiences of Older Mental Health Patients and Their Care Partners Using a Proxy Account to Access Open Notes: Qualitative Interview Study
JMIR Aging 2025;8:e66690
URL: https://aging.jmir.org/2025/1/e66690
doi:10.2196/66690
PMID:

Original Paper

Employers' Perspectives of Caregiver-Friendly Workplace Policies for Caregiver-Employees Caring for Older Adults in Hong Kong: Thematic Analysis

Maggie Man-Sin Lee[1], BBA, MPH, JD, PhD; Eng-Kiong Yeoh[1,2], MBBS; Eliza Lai-Yi Wong[1,2], BSN, RN, MPH, PhD

[1]JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Centre for Health Systems & Policy Research, JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Eliza Lai-Yi Wong, BSN, RN, MPH, PhD
JC School of Public Health and Primary Care
Faculty of Medicine
The Chinese University of Hong Kong
School of Public Health Building, Prince of Wales Hospital
Shatin, New Territories
Hong Kong
China (Hong Kong)
Phone: 852 2252 8772
Fax: 852 2145 7489
Email: lywong@cuhk.edu.hk

Abstract

Background: Caregiver-friendly workplace policies (CFWPs) are rare in Hong Kong. With Hong Kong facing a "silver tsunami" in the near future, it is important to understand the need for such policies and the views of employers for future facilitation.

Objective: This study aimed to identify the support that is currently provided or that could be provided to caregiver-employees (CEs) caring for older adults in Hong Kong and assess the challenge and facilitative support for employers to adopt CFWPs in the specific context of Hong Kong.

Methods: A qualitative research design with semistructured individual in-depth interviews with employers from Hong Kong was adopted for this study. A purposive snowball sampling method was used to recruit participants from the 7 primary industries mentioned in the Hong Kong census and from all 3 employer types (private, public, and nongovernmental organizations), which allowed the inclusion of participants sensitized to the idea and potential of CFWPs. Thematic framework analysis was used to evaluate the data collected during the interviews.

Results: We interviewed 17 employers and managers from 7 major industries in Hong Kong (2.5 to 120,000 employees). There were 4 (24%) male and 13 (76%) female participants, and the participant age ranged from 30 to 50 years. All participants held managerial positions at the time of the interview. Of the 17 participants, 13 were from private companies, 2 were from public institutions, and 2 were from nongovernmental organizations. Four of the companies had a global presence. Four main themes were identified: (1) current support and potential support for CEs (which was limited to discretionary annual leave and unpaid leave when annual leave was exhausted), (2) challenges in adopting CFWPs, (3) facilitating support for adopting CFWPs, and (4) incentives for adopting CFWPs. The participants rated information and resources for CEs (mean 8.56, SD 0.37), bereavement leave (mean 8.47, SD 0.63), flexible working hours (mean 8.32, SD 0.48), and caregiver-inclusive corporate culture (mean 8.32, SD 0.48) as essential CFWPs for CEs in Hong Kong.

Conclusions: While several studies have reported the types of CFWPs and their impacts on CEs, stakeholders' perspectives on CFWPs have been rarely investigated. This study found that although employers consider CFWPs as necessary and see them as a catalyst for a long-term win-win situation, the current support for CEs is discretionary and industry-specific. Government leadership is critical for formulating, piloting, and implementing CFWPs to create a friendly environment that encourages disclosure with trust and respect across industrial sectors in Hong Kong. This study identified the current unmet needs and demands of CEs

from the employer's perspective, the barriers to large-scale adoption of CFWPs, and the path forward to inform further discourse and policy formulation in Hong Kong.

(JMIR Aging 2025;8:e68061) doi:10.2196/68061

KEYWORDS

caregiver; aging; burnout; stress; mental health; employees

Introduction

Filial piety has strong cultural roots in Confucian ideology and is especially prominent in Eastern countries in contrast to children-centered societies, which are more prevalent in Western countries [1-3]. However, there is some evidence of the reduction of the influence of classical Confucian ethos in Asian societies, especially in highly urbanized cities like Beijing [4,5], although the extent of this change is currently unknown. Similarly, the cultural reconfiguration of familial values and expectations has been noted among immigrant Asian families in Western countries [6].

Moreover, emerging research indicates that authoritarian filial piety, which prioritizes social obligations toward parents over one's own needs, varies, but reciprocal filial piety, characterized by genuine gratitude for parents' efforts or sacrifices and voluntary support for them, remains consistent across Eastern and Western cultures [3,7,8]. For instance, a study comparing filial piety among Asian Singaporean and non-Asian Australian adults showed a significantly higher mean score for authoritarian filial piety (mean 21.6, SD 5.9 vs mean 19.4, SD 6.0) on the standardized dual filial piety scale among Singaporean adults, but reciprocal filial piety (mean 34.4, SD 4.8 vs mean 34.2, SD 4.4) was comparable in the 2 populations [8].

Important differences have also been noted in patterns of filial piety within Asian cultures. For example, young adults in Taiwan endorsed strong beliefs in reciprocal and authoritarian filial piety, with both beliefs positively impacting life satisfaction, while young adults in Macau endorsed strong beliefs in reciprocal filial piety and those in Hong Kong endorsed strong beliefs in authoritarian filial piety, with positive impacts on life satisfaction [9]. Interestingly, collectivism and not ethnicity has been identified as a predictor of both forms of filial piety and positive attitudes toward caring for aging parents [7]. Hence, filial piety may be more cross-cultural than previously believed, albeit with regional nuances [7,8]. Indeed, 74% of older adults in the United States National Health and Aging Trends Study (NHATS 2011-2017) reported receiving support from one child caregiver, and 41% of children provided care to aging parents [10].

The rapid increase in the aging population has also increased the caregiving burden of family caregivers, with both being reported across cultures and countries [11,12]. Although several studies have identified caregiver burden as a significant risk factor for burnout and poor mental health outcomes among family caregivers [11,13,14], stronger filial piety may lead to lower caregiver burden [15,16]. A systematic review of 12 studies showed a significant negative correlation (r=–0.23) and association (β=–0.27) between filial piety and caregiver burden among adult children [15]. However, although filial piety is a prominent determinant of caregiver burden, it is not the sole determinant even in Asian societies [17].

While noting mild-to-moderate burden among Asian caregivers of older adults with chronic illnesses, a systematic review observed that specific characteristics of the care recipients (functional dependency, comorbidities, memory, and sleep impairments) and caregivers (advancing age, male gender, spouse as a care recipient, longer care duration, and lack of support/assistance) were associated with higher caregiver burden [16]. However, Asia has a lower old age dependency ratio (population aged ≥65 years to working age population) and health-adjusted dependency ratio (population with the same or higher disease burden to population with a lower disease burden as an average global 65-year-old person) than Europe, North America, or Oceania [18].

East Asia, in particular, has the highest employment-to-population ratio globally (63; for comparison, the corresponding numbers for the Euro area and North America are 58 and 60, respectively) [19], and thus, a substantially large proportion of caregivers in this region likely have a dual role of carer-employee (CE). Studies have shown that those with elderly or child caregiving responsibilities are less favorably viewed by employers in terms of competency (F_2=10.99; P<.001; η^2p=0.092), agency (F_2=15.00; P<.001; η^2p=0.121), commitment (F_2=164.83; P<.001; η^2p=0.602), and availability (F_2=69.01; P<.001; η^2p=0.388) [20]. Hence, CEs were less likely to be hired (F_2=13.65; P<.001; η^2p=0.111) and, if hired, were offered lower salaries (F_2=13.08; P<.001; η^2p=0.107) than nonprimary caregivers [20]. Moreover, the dual role adds extra strain, negatively impacting the mental and physical health of CEs [21,22] and their performance in both roles, which aligns with the role accumulation theory [14,23].

Therefore, the strain of the dual role of CEs will likely offset the low caregiver burden associated with filial piety in Asian populations [8,16,17]. Moreover, the low dependency ratio among Asian populations is likely to be mitigated by the rapid aging of many East Asian populations by 2030, as shown by the China Health and Retirement Longitudinal Study [24]. Thus, several studies have investigated the impact of caregiver-friendly workplace policies (CFWPs) on CEs, which have been associated with the improved overall health of CEs by reducing occupational and overall stress, minimizing work interruptions, and improving performance [22,25-29]. There are also direct economic benefits accruing from adopting CFWPs. For example, educating CEs about their caregiving activities generates a net benefit ranging from US $48,010 to US $675,657 for CEs and employers [30].

CFWPs that help employees manage multiple roles and improve their work-life balance [31] are increasingly being adopted to mitigate some of the caregiving burden of CEs, especially in developed economies [32]. For instance, most employers in the United States and Canada have adopted measures to create a caregiver-friendly workplace environment, such as support services, flexible working hours, financial support, and paid/unpaid leave [33,34]. Similarly, the UK government has set up programs and nongovernmental organizations (NGOs) to help the private sector become more CE friendly [33]. In the specific context of Hong Kong, official caregiver support has not yet been observed, although some private companies provide generic support [35].

Moreover, while several studies have reported the types and impacts of CFWPs [29,33], there is a paucity of studies reporting stakeholders' perspectives on CFWPs. We could identify only 1 study that explored the perspectives of stakeholders (managers working in the Canadian health care sector) on CFWPs [36]. Therefore, this study explored the views and experiences of employers regarding CFWPs for CEs caring for older adults in Hong Kong with the objectives of identifying (1) the support that is currently provided and that could be provided to CEs by employers and (2) the challenge and facilitative support for employers to adopt CFWPs in the specific context of Hong Kong.

Methods

Setting

In Hong Kong, authoritative filial piety emphasizes caring for elderly parents [37]. However, the shift in parental expectations from tangible to emotional support poses challenges for caregiver mental health [38]. Sole children of aging parents face intense burdens, especially in nuclear families, which comprise 64% of households in Hong Kong [39]. One-third of households have at least one elderly member [39], and only 2.3% of the elderly population is entitled to full subsidization [40]. Additionally, given the 56% increase in children with special educational needs in the last half decade [41], support for people in the "sandwich generation" who are caring for both young children and old parents becomes important [28].

The number of CEs, defined as family members (spouse, parents, or siblings) who are employed with monetary rewards and at the same time provide care to another family member [42], in Hong Kong remains undocumented, as general informal caregiver support has only recently started to pick up in Hong Kong. Nevertheless, Hong Kong's highly engaged workforce (50.87% of the overall population) [43] faces additional challenges due to chronic diseases among older adults. Over 65% of the older population in Hong Kong have at least one chronic disease, while one-third have ≥2 chronic diseases [44].

Design

A qualitative research design with semistructured individual in-depth interviews involving employers and company management in Hong Kong was adopted for this study. Independent interviews helped build critical rapport between the researcher and the participants, with minimized privacy issues compared to focus groups, so that respondents would be as open and honest as possible. The COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Multimedia Appendix 1) was used as the reporting framework [45].

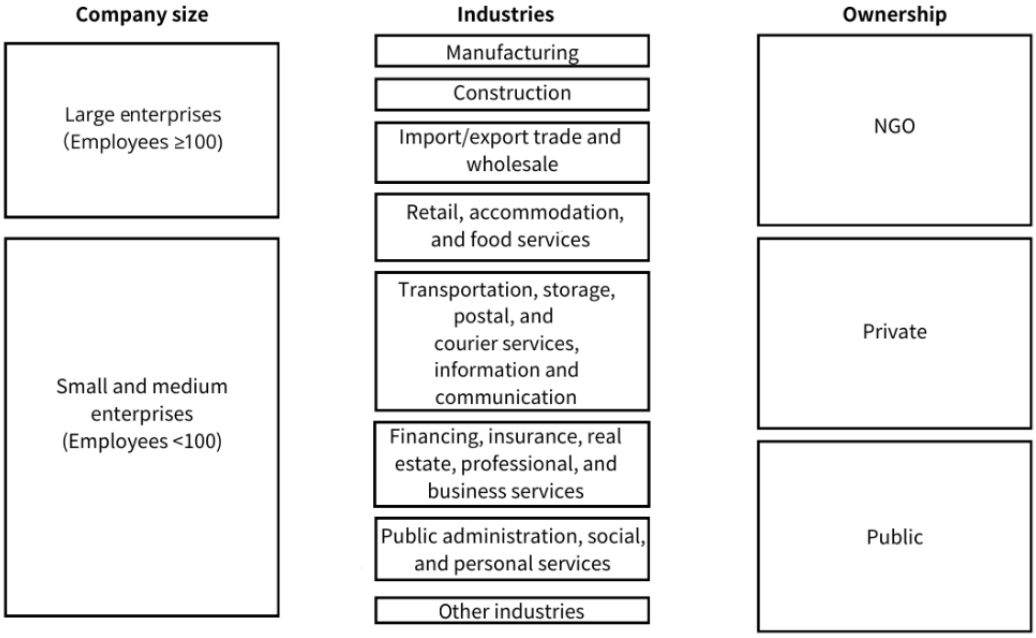
Participants

Purposive sampling was used to recruit participants, which allowed the inclusion of participants sensitized to the idea and potential of CFWPs. The list of potential participants or the sampling frame included individuals who had participated in ESG (environmental, social, and governance) policy consultations, and the information was sourced from a local NGO that advocates for working caregivers. Participants were also identified through the snowballing technique via referrals during interviews. Interviews were arranged until data saturation was achieved.

We started recruitment based on the 7 main industries in the Hong Kong census [39] and recruited from all 3 employer types (private, public, and NGO) (Figure 1). We then expanded the list to ensure a good proportion of large enterprises, small- and medium-sized enterprises, NGOs, and public institutions until data saturation was reached. Data saturation was assumed when no new themes emerged during interviews.

Participants were deemed eligible for inclusion if they had served in a managerial role for at least one staff member at a company, had experience working with CEs in a professional or personal context, and were proficient in English or Chinese (Cantonese or Mandarin). Those serving as third-party consultants for a company were excluded.

Figure 1. Map of employers and management recruited in this study. NGO: nongovernmental organization.



Data Collection

The semistructured interviews were conducted using a discussion guide consisting of 17 questions related to four aspects: (1) respondents' personal experiences with caregiver employees, (2) their attitudes and preferences for policies, (3) their perceptions of a caregiver-friendly workplace, and (4) the rating of 10 policies (Multimedia Appendix 2). The discussion guide was developed based on a thorough literature review, the first author's previous Master's thesis [35], and the opinions of 2 expert panels of researchers specializing in the field of caregiving (Professor E Wong, Professor D Dong, and Dr C Chan from the School of Public Health and Primary Care, The Chinese University of Hong Kong).

The first author (MMSL; PhD candidate at the Chinese University of Hong Kong) conducted all interviews in Cantonese face-to-face at the participants' offices or through videoconferencing, with each lasting 60 to 90 minutes. Participants' written consent and demographic data (age, gender, working position, and marital status) were obtained after the interviewer's self-introduction (comprising qualifications, current and past research focus, and experience) and an explanation of the purpose of conducting the research. Conversation starters were used to open up the participants, so they would feel more at ease and provide the most candid answers. Prompts were also used for specific and related questions during the interview to fully explore their experiences and perceptions [46,47].

At the end of the interview, participants were invited to vote for specific CFWPs based on the degree of importance (0 to 10) from the perspective of employers and managers. The list of CFWPs developed from the global literature review included caregiver-inclusive corporate culture; paid caregiver leave; unpaid caregiver leave; bereavement leave; flexible working hours; flexible work locations; switch to part-time work; unpaid leave; medical needs/insurance of employees' parents; and caregiving information, carer skills, and guide to community care resources. No nonparticipants were present during the interviews.

Analysis

Thematic framework analysis in applied policy research was adopted for data analysis. A 5-stage analytical process described by Ritchie and Spencer [48] was used, which comprised (1) familiarization, (2) identifying a thematic framework, (3) indexing, (4) charting, and (5) mapping and interpretation (Figure 2) [48,49].

Audio recordings of all interviews were transcribed and coded using NVivo software (version 12, Lumivero) independently by 2 researchers (MMSL and EW) to ensure all interpretations were agreeable and to enhance the reliability of the study. The transcripts or the findings were not returned to participants for comment or corrections. Any discrepancy in coding was resolved by discussion with supervisors.

The rating of the 10 policies by importance was reported as mean and SD. Inapplicable policies for specific industries were removed during the calculation. For example, there were no flexible work locations for the retail, food, and construction industries. A coding hierarchy chart (Figure 3) was generated to show the visual patterns of the coding.

Figure 2. The 5-stage analytical process used in this study.

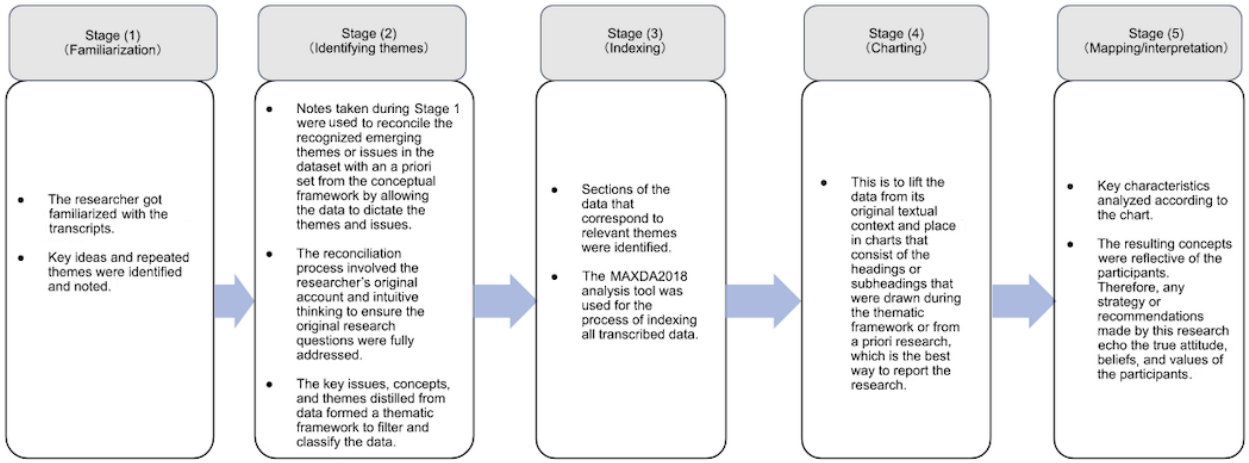
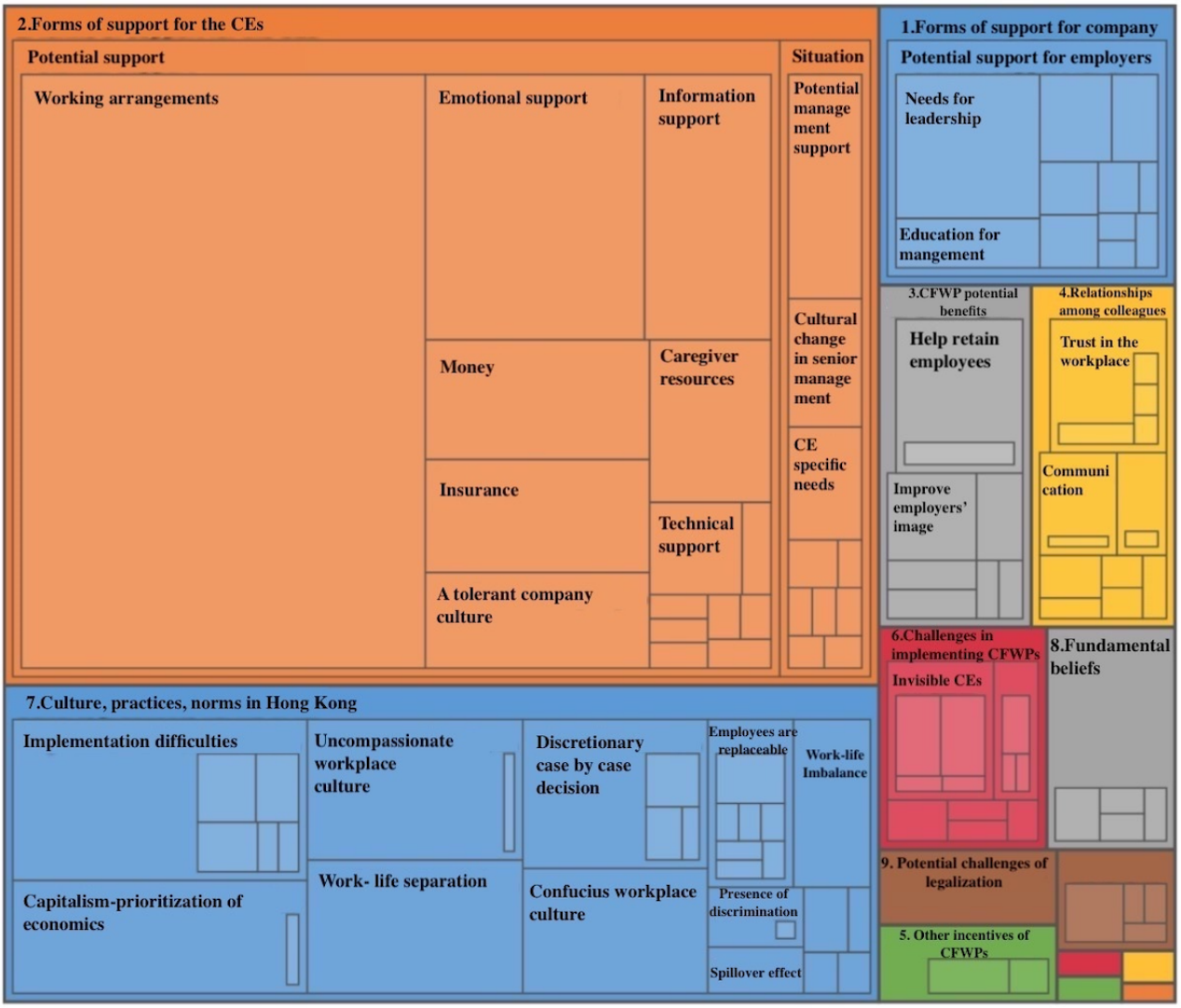


Figure 3. Coding hierarchy chart. CE: caregiver-employee; CFWP: caregiver-friendly workplace policy.



Ethical Considerations

Research ethics approval was granted by The Chinese University of HK Survey and Behavioral Research Ethics Committee (reference number: SBRE(R)-21-011). Written consent was obtained from interview participants, and they were assured of their rights and freedom to withdraw from the study at any time. The information and responses of the participants were treated confidentially. All project data were anonymized and kept in

password-protected folders that were only accessible to the project researchers and supervisors.

Results

Overview

Interviews with 17 participants comprising 4 (24%) male and 13 (76%) female participants were conducted in November 2021, and additional participants were not recruited as data saturation was achieved. Details of the participants are presented in Table 1. The age of the participants ranged from 30 to 50 years, and all held managerial positions at the time of the interview. Of the 17 participants, 13 were from private companies, 2 were from public institutions, and 2 were from NGOs. Given that the proportion of workers employed in public institutions and NGOs is less than 25% of the total workforce in Hong Kong [39], the ratio of 13:2:2 is justified.

The companies and institutions encompassed the 7 major industries in Hong Kong (ranging from 2.5 to 120,000 employees). Four of the companies had a global presence. Data saturation was achieved for NGOs and public institutions with 4 participants, all of whom viewed their industries' organizational policies and culture similarly.

Most participants (16/17) expressed that the concepts of CFWPs had been unheard of before the interviews. However, they all supported the core ideas of CFWPs (enabling the dual roles of CEs as a caregiver and a productive member of the institution) and agreed on the fact that CFWPs were needed, subject to the specific needs of the business, for the betterment of the venture and Hong Kong's future. The participants rated (1) information, caregiving skills, and guide to community care resources for CEs; (2) bereavement leave; (3) flexible working hours; and (4) caregiver-inclusive corporate culture as the most important policies (Table 2).

Despite the unanimity in supporting CFWPs, the operational practicality of CFWPs was of concern to all the participants. Four overarching themes were identified as outcomes of the thematic framework method in the analysis: (1) current support and potential support for CEs, (2) challenges in adopting CFWPs, (3) facilitating support for adopting CFWPs, and (4) incentives for adopting CFWPs (Table 3).

Table 1. Detailed demographics of the participants.

Number	Organization				Position	Age (years)/sex	Marital status[a]
	Industry	Type	Size	Presence			
1	Public administration, social, and personal services	Public	100 (LE[b])	Regional	Chief/director of a primary school	47/male	Married; #2
2	Public administration, social, and personal services	Private	8-10 (SME[c])	Regional	Executive director of a beauty company	60/female	Married; #1
3	Transport, storage, postal, and courier services; information and communication	Private	8-10 (SME)	Regional	Director of an e-platform building firm	40-50/female	Separated; #0
4	Import/export and wholesale	Private	8-10 (SME)	Regional	Director of an import/export company	30-40/female	Married; #1
5	Financing, insurance, real estate, professional, and business services	Private	200-300 (LE)	Regional	Managing director of a financial company	36/female	Single; #0
6	Public administration, social, and personal services	Private	9000 (LE)	Global	Senior marketing director of a music label corporation	48/female	Married; #2
7	Public administration, social, and personal services	Private	2.5 (SME)	Regional	Director of an individual psychiatric clinic	40/male	Married; #1
8	Retail, accommodation, and food services	Private	32 (SME)	Regional	Sales and marketing director	32/female	Single; #0
9	Public administration, social, and personal services	NGO[d]	15 (SME)	Regional	Project manager	35/female	Married; #0
10	Financing, insurance, real estate, professional, and business services	Private	20,000 (LE)	Global	District director	60/male	Married; #2
11	Public administration, social, and personal services	Public	1300 (LE)	Regional	Head of department at a financial regulatory institution	47/female	Married; #2
12	Public administration, social, and personal services	NGO	40 (SME)	Regional	Manager of a district health center	30-40/female	Married; #0
13	Financing, insurance, real estate, professional, and business services	Private	100 (LE)	Regional	Senior marketing manager of a marketing company	32/female	Single; #0
14	Public administration, social, and personal services	Private	120,000 (LE)	Global	Market access manager of a pharmaceutical company	30-40/female	Single; #0
15	Manufacturing	Private	100 (LE)	Global	CEO[e]	30-40/male	Married; #2
16	Construction	Private	10 (SME)	Regional	Owner of a construction company	31/female	Married; #1
17	Retail, accommodation, and food services	Private	11 (SME)	Regional	Owner of a restaurant	31/female	Married; #0

[a]# denotes the number of children.

[b]LE: large enterprise.

[c]SME: small and medium enterprise.

[d]NGO: nongovernmental organization.

[e]CEO: chief executive officer.

Table 2. Participant rating of the importance of the policies.

Rank	Policy	Score, mean (SD)
1	Information/caregiving skills/guide to community care resources	8.56 (0.37)
2	Bereavement leave	8.50 (0.55)
3	Flexible working hours	8.47 (0.63)
4	Caregiver-inclusive corporate culture	8.32 (0.48)
5	Flexible work locations	7.80 (0.61)
6	Paid caregiver leave (especially for those caring for elderly people with chronic diseases, excluding counting maternity leave)	7.72 (0.54)
7	Unpaid caregiver leave	7.60 (0.62)
8	Aiding medical needs/insurance of employees' parents	7.56 (0.75)
9	Switch to part-time work	7.03 (0.44)
10	Unpaid leave	6.78 (0.68)

Table 3. Themes and subthemes.

Main theme	Subthemes
Current support and potential support for CEs[a]	• Leave • Nonfinancial support • Emotional support • Information support • Technical support • Industry-specific support
Challenges in adopting CFWPs[b]	• Invisibility of CEs • Hard to adopt a fair policy for the duration of leave • Concerns about standardization
Facilitating support for adopting CFWPs	• Establishment of a government compensation fund and legal framework for caregiver leave • Government leadership
Incentives for adopting CFWPs	• Build a positive organizational culture • Recruitment of talent and extension of the worldwide network

[a]CE: carer-employee.

[b]CFWP: caregiver-friendly workplace policy.

Theme 1: Current Support and Potential Support for CEs

Six subthemes were identified based on the participants' current and potential support needs.

Leave

All participants had access to discretionary annual leave and unpaid leave when annual leave was exhausted, except participants 16 and 17, who were from the construction and food industries, which grant pay on a daily or hourly basis. All participants stated that they had the flexibility of taking a few hours of leave during working hours at their discretion for an emergency, for instance, a medical emergency related to an older family member at home. However, none of the participants reported specific caregiver policies in their workplaces.

There was bereavement leave in the participants' companies, but as noted by participant 1, such leave "is only for immediate family members." Participant 1 further noted the rigidity of bereavement leave as follows:

> *Like my wife's father, whom she also cared for, was paralyzed for a while and then passed away. I could not take leave for his funeral.*

Nonfinancial Support

Unresolved age-related care challenges created pessimism about the future among 10 participants: rising health care costs and longer life expectancy strain lower- and middle-class families, and single-child families bear immense burdens when caring for elderly people and young people. Without the financial means to hire a helper or buy insurance policies, the quality of life of CEs and their families would drastically drop in the face of immense economic, physical, mental, and emotional strains. Employers recognized the resource-based challenges CEs face, emphasized the need for nonfinancial support, and deemed direct financial support too expensive to be on the agenda for the time being. Participant 4 specifically pointed to the

importance of time as an example of nonfinancial support for working caregivers:

> *I think the most important thing is time. There are different degrees of illness; maybe he (i.e., the care recipient) has a stroke, or he has terminal cancer, a case like mine has only a few months left. And if you have to keep working and cannot spend time with your family. I think that's a great pity.*

Emotional Support

Seven participants pointed to the importance of emotional support for CEs. Participants pointed to the heavy role burdens of CEs, who need to fulfill their roles as parents, team members in the workplace, and caregivers of older adults. For instance, participant 8 witnessed severe psychological distress, possibly an emotional breakdown, during one of the work trips when her colleague was struggling with poor support from her marriage and was struggling in her career while caring for her children and an ailing elderly person. Participant 4 considered that emotional support should apply to all employees (not only CEs):

> *Emotional support should not only apply to CEs. I think the emotions of all employees are important as long as they are in the workforce. Emotional distress is like a cold, very common.*

Participant 12, an occupational therapist specializing in mental health, noted that work stress in Hong Kong is so severe that depression in young workers is often due to work stress. Long working hours and the stigma and discrimination against mental illness would further discourage Hong Kong CEs from taking care of their mental well-being. A sudden imposition of the caregiver role would further challenge the coping skills of the workers themselves and jeopardize their well-being.

Information Support

Four participants pointed out the need for a hotline for CEs to receive information and counseling services. Participant 14, who worked for a multinational pharmaceutical company, reported that they had already set up an international hotline to support employees in potentially all life-related aspects:

> *You can call there to find out about the best schools for children in your neighborhood. Since the hotline is run by a third party, anonymity is maintained to protect privacy, which makes the service even more attractive.*

Technical Support

Four participants also pointed out the problem of time strain (high time demands and little control over timing), which applies to both work and unremitting care work. They therefore pointed to the importance of home-based care services that allow CEs some personal space and time. In this context, participant 6 mentioned that today's open-plan offices prevent CEs from taking personal calls when needed. Therefore, personal space in the workplace, such as a phone box, would support CEs in fulfilling both roles effectively and safely. Participant 6 elaborated as follows:

> *Open-plan offices are very popular today, which I object to because it lacks personal space. A phone box is a personal space where colleagues have to take their own personal calls. For example, if I were to have a conversation with my mother, who is also deaf. I yell, 'No, the diarrhea pill is upstairs.' That can be very embarrassing in an open-plan office. In fact, I think this concept is in line with the idea of work-life separation.*

Industry-Specific Support

Participant 1, one of the directors at a publicly funded primary school, pointed out that staff turnover could severely affect the quality of teaching and that CFWPs were crucial to the school context, given the overly protected privacy in the school's workplace culture. No personal matters were allowed to be revealed that could cause suspicion and doubt among colleagues and between teachers and parents. According to participant 1, in this specific sector, either flexible working hours and locations or a switch to part-time work was possible. Given the bureaucracy in applying for leave and extended working hours, a CFWP culture for school teachers was necessary. Participant 1 commented:

> *Changes should be made at the top policy level of school administration. Because at the moment, even if you are a headmaster, you act according to a set of rules that restrict what you can do and what you can use.*

How wages are paid is also a determinant of whether a policy of flexible working hours is needed. In construction (daily rates), insurance (piece rates), or creative industries (commission rates), the flexibility of working hours is built into the mode of payment. Therefore, there is no substantial need to include it in CFWPs. Flexible work location arrangements are not applicable in construction, retail, food, and personal services where the location of the staff is essential to the purpose of the employment.

Theme 2: Challenges in Adopting CFWPs

Invisibility of CEs

Many challenges of introducing CFWPs are closely linked to the nature of CEs. A key feature of CEs is their invisibility. Unlike maternity leave, which 11 participants said is easier to implement than caregiver leave, no identifiable physical features (such as a swollen and protruding belly) make CEs recognizable. Participant 14, for instance, made the following statement:

> *People are very visual, maternity leave, that the colleague just had a baby. But caregivers, not everyone can see their parents or the needs of their family.*

The invisibility of CEs has enormous implications for the adoption and implementation of CFWPs, as it becomes more difficult to convince everyone of the need to support CEs and make extra efforts to accommodate them, even if maternity, paternity, and caregiving leaves are all based on the same basic need to care for family members. Participant 17 elaborated as follows:

> *[regarding employees requiring maternity leave], we can see the belly sticking out. Also, there is a life in there; you are responsible if something happens to the mother or child. You do not feel so strong if the colleague has a sick parent, for example. If the pregnant colleague feels uncomfortable with her belly and lies down, you have to let her immediately. But if the colleague says that her father is in hospital because of stomach problems, you would not feel so strongly at that moment. Besides, I have experienced that myself—some people might doubt if it is even real.*

According to the participants, CEs disclose their status as caregivers directly to their supervisors when they need to take leave or use other forms of support.

Hard to Adopt a Fair Policy for the Duration of Leave

The ambiguous needs of CEs create many grey areas for institutional arrangements. To be fair to all employees, the rationale for the arrangements should be clear and relevant, but this poses a problem for participants when it is necessary to allocate a certain number of leave days depending on the nature, severity, and stage of medical conditions. On the other hand, participants lacked the medical expertise to assess the CE's situation for a case-by-case scenario, while adopting an all-caregiver-inclusive package for all CEs could be resource-intensive and disruptive to daily operations. Thus, there was no consensus among the participants regarding standardization. Participants 9 and 11, for example, felt that standardization would provide a reference point and guideline for companies, which would help combat discrimination against CEs and give the company an advantage in hiring workers. In contrast, others, such as participant 6, felt the standardization would lack the flexibility to be helpful. This complex and difficult-to-standardize situation of CEs contrasts with maternity and paternity leaves, which are standardized and predictable.

This ambiguous need of CEs further discourages companies from using CFWPs when they entail prohibitive costs in the long run. Participant 15 explained:

> *There is a predictable time frame for maternity and paternity leave... But sickness is a bit longer. For maternity or paternity leave, you would give birth to maybe only two to three children in this lifetime and then stop, so you would only take maternity/paternity leave two to three times; here, there is a fixed number of dates, you can predict it. But if you are taking care of chronically ill or dependent family members, it may take a long time; it may be 5 or 10 years. In between, you may have to take leave every time to accompany them to doctor's appointments, which may happen 10-20 times a year; that possibility actually exists. So, with that uncertainty, it becomes a bit more difficult to implement* [the CFWP policy]

Concerns About Standardization

Standardization also raises concerns that the mechanisms of CFWPs could be abused. Participant 16 pointed to the difficulty of proving caregiver status:

> *Can I just find any aunt or distant relative? That is, the moment I want to take a holiday, the moment I do not want to take this job, I can just find a random aunt. They give me a paper, I give it to the company, then I get the benefits. I feel that these grey areas can be abused very easily... If I introduce care leave as an employer, I may suddenly get them [these medical certificates]. The next day, I wake up and suddenly get a message that I can not go to work [because of a family member's condition]... The wheel would stop turning; we would not know what to do, so I think this thing is more difficult than paternity leave. With paternity leave, you can easily prove that you are the father. Due to the very ambiguous nature of CEs' needs, this can be a major trust issue for management to support CEs.*

Participants commonly expressed that it was difficult to justify the CFWPs giving priority exclusively to workers caring for older individuals, risking favoritism over other needs, such as workers raising their children. The heavy workload of CEs for ailing older individuals would not make caring for needy young individuals less stressful. The cultural preference for newborns in Hong Kong was pointed out by participant 4 as follows:

> *Everyone expects a happy baby, but not everyone would consider the possibility of having a sick parent.*

The comments make the different treatment of CEs with an ailing older relative even more challenging to accept.

Theme 3: Facilitating Support for Adopting CFWPs

Governmental Compensation Fund and Legal Framework for Caregiver Leave

Addressing paid/unpaid leave for caregivers involves considering financial costs, a central issue raised by all participants. Participants proposed government compensation funds for employers to cover the costs of CEs' time off, cashing out the mandatory provident fund (MPF) during a medical crisis, legal frameworks for eligibility, and compulsory employment insurance. Regarding the cash out from the MPF, participant 17 commented:

> *The original intention of the MPF was to maintain a certain quality of life in retirement. But if he (CE) can't get through the hardest time of his life (caring for ailing older family members), it doesn't matter how much money he'll receive in the future.*

Government Leadership

The practical challenges practicing doctors face include a lack of government guidance and reimbursement. Many participants felt that the government must take the initiative to support CEs. Many participants expressed the opinion that "if the government implements it first, we will follow;" however, there is currently no single, consistent CFWP in Hong Kong. Therefore, support for CEs has only been granted on a discretionary and case-by-case basis. If enacted, the policy and legislation supporting CEs would be mandatory for them to follow, and this government initiative would even trigger a shift in

companies and management thinking. Participant 2, for example, said the following:

> *I do not think there was anything like maternity leave before; it's even more similar with paternity, but we all have that now, so I would think, why not? [I am] also very, very supportive* [of CE leave]

Two participants also felt that with the introduction of maternity and paternity leaves, it would be easier to advocate for support for caregivers if the government takes the lead in implementing policies. Participant 2 commented as follows:

> *There is paternity leave now. Before, it was difficult, but now it is possible. If this was applied to CEs, maybe it would be quicker [to get support for them] because people have already taken a step forward.*

This was also affirmed by participants 3 and 8, showing some level of readiness for CFWPs in Hong Kong, but it is crucial for the government to take the lead in getting this process started.

In the context of introducing a legal framework to support the future implementation of CFWPs, participant 11, a department head of a financial regulatory institution, highlighted the earlier success of the pilot project on maternity leave before it was legalized for all companies:

> *The Government could test the leave program for CE and other CFWP in some selected departments to identify the need and the possibilities of implementation. After consultations with various sectors of the economy, I believe it would be readily adopted.*

All participants expressed willingness to introduce some forms of CFWPs if the government could first demonstrate the benefits of the measures and set the framework for practice. Participant 13 felt that a pilot phase within the government is a guarantee for the private sector and is one of the many conditions for local businesses to consider CFWPs.

Four participants were aware of the importance of the government communicating the needs of CFWPs to the business sector, particularly for management understanding, including recognition of the role of CEs, potential caregiver burnout, disease progression, management of older people, and stigma/discrimination against CEs. Participant 6, a director at a multinational company, emphasized that the human resources department could take the lead in advocating for CFWPs in the cultural and operational contexts to ensure sustainability and enforcement. Participants 4 and 6 recommended using case studies by the government or those in business to explore the positive changes associated with CFWPs and motivate other companies to create a caregiver-friendly workplace.

Theme 4: Incentives for Adopting CFWPs

Build a Positive Organizational Culture

The improvement in loyalty and sense of belonging that CFWPs bring in the long term was acknowledged. Participant 10, a district director of a financial institution, noted the following:

> *supporting CEs will result in productivity losses—as an immediate loss to the company, but in the long run, the benefits of improving loyalty and sense of belonging would outweigh the costs overall.*

Participants 2 and 5 noted positive spillover effects, specifically, work engagement improving the caregiving experience. CFWPs aim to facilitate a balance between the dual roles, allowing for healthier management of CEs' circumstances, which is an incentive for adoption. Participant 5 commented:

> *It's better to stay at work than to stop for caregiving because you still get to see what's going on in the world, and your brain keeps moving.*

Round-the-clock caregiving duties could be more stressful than taking on dual roles.

CFWPs could be seen as a confidence-building exercise between the management and employees. Management's effort to cultivate a caregiver-friendly workplace culture is a prerequisite for gaining CEs' trust before disclosing their status and asking for consideration for their individual needs. Participant 6 referred to her experience as follows:

> *When my employee came to my office and asked for a day off to take care of his parents, he must have trusted me that he was asking for help. I am very grateful for that trust.*

To strengthen trust, participant 4 suggested using a medical certificate (similar to a carer's passport scheme in the United Kingdom [50]) confirming the caregiving role of the employee, with constant updates on the care recipient's illness, and indications of possible commitments and the durations of such commitments. Participant 3 was aware of the guilt of CEs about not being able to fulfill the dual roles simultaneously. This sense of guilt could not only hinder productivity but also encourage the suppression of emotions and interfere with communication.

Recruitment of Talent and Extension of the Worldwide Network

Most participants explicitly expressed their belief that CFWPs could give companies an advantage in recruiting talent. Participants considered that CFWPs allow for the retention of highly qualified staff and thus more sustainable productivity, as well as a potential reduction in training and recruitment costs. However, participant 12 stressed that the attractiveness of CFWPs might be limited among younger workers.

Participant 14 emphasized that if nonfinancial reporting is considered, CFWPs could make a company a worthwhile investment for potential investors under ESG standards:

> *Impact investing is now becoming more popular. Many global companies have already included ESG in their reporting; we are already lagging behind a lot, although our company has already set up a hotline for general employee concerns.*

For Hong Kong, there are overall benefits of companies taking up CFWPs, namely keeping workers employable rather than out of the workforce for a significant duration, attracting foreign talent to Hong Kong, and reducing public health costs when CEs perform caregiving duties more efficiently and effectively.

Discussion

Principal Findings

Although support for CEs in Hong Kong is currently discretionary and depends on the context of industries, this study identified strong motivations for Hong Kong employers to adopt CFWPs. While improving mental well-being can lead to a direct improvement in productivity [51-54], the benefits of CFWPs may be more extensive than anticipated. Beyond indirect operational incentives, such as employee loyalty and talent retention, CFWPs could make a company an attractive investment for socially conscious investors under ESG standards [55]. Furthermore, CFWPs can have a positive spillover effect with a potentially significant impact on CEs' mental well-being, especially when caregiver-inclusive corporate culture is more explicit than making decisions on a case-by-case basis.

Although this study identified several gaps and barriers to the large-scale adoption of CFWPs in Hong Kong, there is ample precedence of organizational changes in Hong Kong's public and private sectors. For instance, the 8-step theory of change by Kotter [56], which involves the following steps: (1) create urgency, (2) build a guiding coalition, (3) form a strategic vision, (4) communicate the vision, (5) empower others to act on the vision, (6) generate quick wins, (7) sustain momentum, and (8) institutionalize new approaches, has been shown to be applicable across Hong Kong's monolithic civil service organizations [57], publicly listed corporations [58], and NGOs [59,60], which represent the 3 primary stakeholders for the large-scale adoption of CFWPs in Hong Kong.

In line with the theory by Kotter [56], employers and management participants in this study echoed the prerequisites for the government to take the leadership in raising awareness of the needs of CEs (step 1: create urgency) and advance the discussion on CFWPs among stakeholders (step 2: build a guiding coalition) to develop (step 3: form a strategic vision) and implement (step 4: communicate the vision) goal-oriented strategies. Government leadership is indispensable for small and medium enterprises as, unlike large enterprises, they lack the in-house resources and experiences to formulate relevant policies and account for more than 98% of businesses in Hong Kong [61] (step 5: empower others).

The piloting of CFWPs in the management of public institutions seems to be an appropriate starting point, as this is where the bureaucratic burden is most remarkable (step 6: generate quick wins). CFWPs could be better thought through in this complex environment before being rolled out to the private sector and NGOs. Notably, the government has recently launched a 2-year pilot program to subsidize caregivers from low-income families caring for older individuals with moderate and severe impairments (at least 3 hours of care per day) to the tune of HK $3000 (US $386) per care recipient [62]. However, the number of households benefiting from such high thresholds is questionable, and the policy does not extend to caregivers who continue to work and contribute to the economy.

Furthermore, our participants indicated the need for a general CFWP with some specific policies for certain industries (step 7: sustain momentum). For instance, information, caregiving skills, and guide to community care resources; bereavement leave; and caregiver-inclusive corporate culture can apply to all industries. With Hong Kong well-positioned as a knowledge-based economy [63], these CFWPs may be immediately relevant for human capital and knowledge-intensive industries, such as social and personal services, professional and business services, financing, insurance, and real estate. However, flexible working hours or locations depend on industry-specific payment arrangements and the individual requirements of the working environment. For instance, construction, retail, accommodation, and food service sectors may require special considerations due to their demographics (eg, relatively younger workers) and wage payment mode (eg, piecework in construction and hourly in food services).

According to the rated importance, providing information resources, cultivating a caregiver-inclusive corporate culture, and introducing flexible working hours and locations could be piloted within the government structure to provide feedback for future public implementation (step 8: institutionalize new approaches). In addition, legalizing paid or unpaid caregiver leave is necessary. While the government reimburses companies for maternity and paternity leaves, there are other options to facilitate the future legalization of caregiver leave, such as cashing out from the MPF and introducing a legal framework on the qualifications and durations of leaves for CEs. However, cashing out from the MPF could jeopardize the original intention of securing financial security in retirement. Thus, a greater government-industry engagement focused on CFWPs is essential.

Our study underscores the need for consensus among stakeholders on caregiver eligibility criteria, including the care recipient's age, care forms, employment conditions, relationships, and medical needs. Quantifiable parameters can ensure fairness and prevent policy misuse. For instance, the voluntary leave-sharing program in the United States allows workers to donate their unused annual leaves directly to another worker who is experiencing a personal or family medical emergency and has exhausted available paid leaves [64,65]. However, the program requires a significant income loss (at least 24 hours for full-time employees) due to illness after exhausting paid leaves [64,65]. Such a program with objective and fair parameters may be considered for Hong Kong's different industrial sectors. Furthermore, such programs may be instrumental in breaking the status quo that personal or family matters do not belong in the workplace and may enable generous interworker benefit transfers.

There is some international precedence in identifying the pathway to develop and implement CFWPs in Hong Kong. For instance, the advocacy of CFWPs is typically supported by NGOs working with governments and legislatures to inform and influence policy [66-68]. While such advocacy groups do not exist in Hong Kong yet, they can be a vital conduit for engaging with CEs and negotiating with the government to address their most pressing needs. Similarly, Canada offers caregivers the world's longest unpaid compassionate leave (26 weeks) [69,70]. Caregiver leave to attend to an ailing parent is covered by compulsory employment insurance, which is funded

by worker premiums, without burdening businesses directly [70]. Similarly, the Japanese government has subsidized 93 caregiver leave days for the business sector [71].

Notably, the legalization of caregiver leave in Taiwan and Japan was in response to an alarming number of caregivers committing suicide or murdering elderly members under overwhelming "caring fatigue" [72]. Hong Kong should urgently consider CFWPs before caring for the needs of the aging population takes too heavy a toll on the working population. Although successful precedents in other countries can be a good starting point, it is essential to consider the unique needs and expectations of CEs and their employers in Hong Kong. For instance, while unpaid caregiver leave is legalized for CEs in Taiwan and Japan [68,71], employers in this study emphasized the importance of paid leave over unpaid leave in line with maternity and paternity leaves in Hong Kong.

Limitations

Our study has several limitations that must be noted. First, our participants were predominantly recruited using the chain-referral method, which is known to introduce selection bias [73]. Thus, the findings may have lower generalizability, representativeness, or external validity than research conducted with a random sample [73]. However, our goal was to identify perceived barriers to CFWP implementation and galvanize further research in this domain, which may, in turn, identify additional barriers. Thus, in this type of research, the inherent risk of selection bias due to chain-referral sampling is not expected to skew the findings significantly. Second, some industries may be underrepresented in our study. For instance, while data saturation was achieved with at least two participants for each category by size (ie, small and medium enterprises and large corporations) and constitution (ie, NGOs, public institutions, and private companies), there was at least one participant from each of the main industrial sectors in Hong Kong. Third, there was no "third party" interviewer to reduce the impact of the researchers' biases and beliefs [74] about the phenomenon (ie, CFWPs) on the quality of the data collected. Finally, this research was conducted at the peak of the COVID-19 pandemic, which could have influenced the participants' perspectives and positions on CFWPs [75]. In addition, the pandemic conditions may have led to the unavailability of participants with diverse opinions [75].

Conclusion

Employers and management recognize the importance of CFWPs, but current support for caregivers is discretionary and industry-specific. Government leadership is crucial for the future implementation of CFWPs, and piloting such policies within the government can drive institutional change. Ongoing consultation with the business sector on the specific needs of particular industries is essential before legalizing unpaid or paid caregiver leave in Hong Kong.

Acknowledgments

We are grateful to MOV Data Collection Center Limited for assistance with the recruitment and data collection phase. This study was funded by The Chinese University of Hong Kong and the Carers Hong Kong Company Limited, a licensed charity in Hong Kong.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[PDF File (Adobe PDF File), 372 KB - aging_v8i1e68061_app1.pdf]

Multimedia Appendix 2
Discussion guide.
[DOCX File , 19 KB - aging_v8i1e68061_app2.docx]

References

1. Ni C. A Comparative Study of Filial Piety and Endowment Patterns Between China and the West. In: Proceedings of the 3rd International Conference on Advances in Management Science and Engineering. 2020 Presented at: 3rd International Conference on Advances in Management Science and Engineering; January 12-13, 2020; Wuhan, China. [doi: 10.2991/aebmr.k.200402.023]
2. Yuan X, Wang Q. A tentative study on differences and integration of Sino-Western filial piety culture. Asian Social Science 2011 Jul 31;7(8):97. [doi: 10.5539/ass.v7n8p97]
3. Gallois C, Giles H, Ota H, Pierson H, Ng SH, Lim TS, et al. Intergenerational Communication across the Pacific Rim: The Impact of Filial Piety. In: Adair JG, Dion KL, Lasry JC, editors. Latest Contributions to Cross-cultural Psychology. England, UK: Routledge; 2020:192-211.
4. Sappor G. A comparative study of filial piety in Beijing and the UK. Journal of Education and Culture Studies 2021 Jan 25;5(1):p22. [doi: 10.22158/jecs.v5n1p22]

5. Lin L, Wang Q. Adolescents' filial piety attitudes in relation to their perceived parenting styles: an urban-rural comparative longitudinal study in China. Front Psychol 2021;12:750751 [FREE Full text] [doi: 10.3389/fpsyg.2021.750751] [Medline: 35140649]
6. Montayre J, Saravanakumar P, Zhao I, Holroyd E, Adams J, Neville S. Holding on and letting go: Views about filial piety among adult children living in New Zealand. J Clin Nurs 2022 Oct;31(19-20):2797-2804. [doi: 10.1111/jocn.16098] [Medline: 34687112]
7. Lim AJ, Lau CYH, Cheng C. Applying the dual filial piety model in the United States: a comparison of filial piety between Asian Americans and Caucasian Americans. Front Psychol 2021 Feb 3;12:786609 [FREE Full text] [doi: 10.3389/fpsyg.2021.786609] [Medline: 35185688]
8. Li WW, Singh S, Keerthigha C. A cross-cultural study of filial piety and palliative care knowledge: moderating effect of culture and universality of filial piety. Front Psychol 2021;12:787724 [FREE Full text] [doi: 10.3389/fpsyg.2021.787724] [Medline: 34925189]
9. Chen WW, Wu CW, Zhou Y, Lo YJ. A within-cultural comparison of filial piety beliefs among college students in Hong Kong, Taiwan and Macau. Families, Relationships and Societies 2020;9(3):361-381. [doi: 10.1332/204674319X15615536876902]
10. Lin I, Wolf DA. Division of parent care among adult children. J Gerontol B Psychol Sci Soc Sci 2020 Nov 13;75(10):2230-2239 [FREE Full text] [doi: 10.1093/geronb/gbz162] [Medline: 31883022]
11. Choi JY, Lee SH, Yu S. Exploring factors influencing caregiver burden: a systematic review of family caregivers of older adults with chronic illness in local communities. Healthcare (Basel) 2024 May 13;12(10):1002 [FREE Full text] [doi: 10.3390/healthcare12101002] [Medline: 38786412]
12. Chang AY, Skirbekk VF, Tyrovolas S, Kassebaum NJ, Dieleman JL. Measuring population ageing: an analysis of the Global Burden of Disease Study 2017. Lancet Public Health 2019 Mar;4(3):e159-e167 [FREE Full text] [doi: 10.1016/S2468-2667(19)30019-2] [Medline: 30851869]
13. Del-Pino-Casado R, Priego-Cubero E, López-Martínez C, Orgeta V. Subjective caregiver burden and anxiety in informal caregivers: A systematic review and meta-analysis. PLoS One 2021;16(3):e0247143 [FREE Full text] [doi: 10.1371/journal.pone.0247143] [Medline: 33647035]
14. Gérain P, Zech E. Do informal caregivers experience more burnout? A meta-analytic study. Psychol Health Med 2021 Feb;26(2):145-161. [doi: 10.1080/13548506.2020.1803372] [Medline: 32816530]
15. Pan Y, Chen R, Yang D. The relationship between filial piety and caregiver burden among adult children: A systematic review and meta-analysis. Geriatr Nurs 2022;43:113-123 [FREE Full text] [doi: 10.1016/j.gerinurse.2021.10.024] [Medline: 34864295]
16. Isac C, Lee P, Arulappan J. Older adults with chronic illness - Caregiver burden in the Asian context: A systematic review. Patient Educ Couns 2021 Dec;104(12):2912-2921. [doi: 10.1016/j.pec.2021.04.021] [Medline: 33958255]
17. Zhang L, Han Y, Ma Y, Xu Z, Fang Y. Eastern perspectives on roles, responsibilities and filial piety: A case study. Nurs Ethics 2021 May;28(3):327-345. [doi: 10.1177/0969733020934143] [Medline: 32666888]
18. Skirbekk V, Dieleman JL, Stonawski M, Fejkiel K, Tyrovolas S, Chang AY. The health-adjusted dependency ratio as a new global measure of the burden of ageing: a population-based study. Lancet Healthy Longev 2022 May;3(5):e332-e338 [FREE Full text] [doi: 10.1016/S2666-7568(22)00075-7] [Medline: 36098308]
19. ILO Modelled Estimates and Projections database (ILOEST). World Bank. URL: https://data.worldbank.org/indicator/SL.EMP.TOTL.SP.ZS [accessed 2025-03-21]
20. Henle CA, Fisher GG, McCarthy J, Prince MA, Mattingly VP, Clancy RL. Eldercare and childcare: how does caregiving responsibility affect job discrimination? J Bus Psychol 2019 Mar 7;35(1):59-83. [doi: 10.1007/s10869-019-09618-x]
21. Gaugler JE, Pestka DL, Davila H, Sales R, Owen G, Baumgartner SA, et al. The complexities of family caregiving at work: a mixed-methods study. Int J Aging Hum Dev 2018 Dec;87(4):347-376 [FREE Full text] [doi: 10.1177/0091415017752936] [Medline: 29345147]
22. Yang J, Wang L, Williams A. The role of carer-friendly workplace policies and social support in relation to the mental health of carer-employees. Health & Social Care in the Community 2023 Mar 13;2023:1-11. [doi: 10.1155/2023/5749542]
23. Kayaalp A, Page KJ, Rospenda KM. Caregiver burden, work-family conflict, family-work conflict, and mental health of caregivers: a mediational longitudinal study. Work Stress 2021 Oct 14;35(3):217-240 [FREE Full text] [doi: 10.1080/02678373.2020.1832609] [Medline: 34483432]
24. Gong J, Wang G, Wang Y, Chen X, Chen Y, Meng Q, et al. Nowcasting and forecasting the care needs of the older population in China: analysis of data from the China Health and Retirement Longitudinal Study (CHARLS). The Lancet Public Health 2022 Dec;7(12):e1005-e1013. [doi: 10.1016/s2468-2667(22)00203-1]
25. Kim J, Ingersoll-Dayton B, Kwak M. Balancing eldercare and employment: the role of work interruptions and supportive employers. J Appl Gerontol 2013 Apr;32(3):347-369. [doi: 10.1177/0733464811423647] [Medline: 25474393]
26. Ding R, Dardas A, Wang L, Williams A. Evaluation of a caregiver-friendly workplace program intervention on the health of full-time caregiver employees: a time series analysis of intervention effects. J Occup Environ Med 2020 Oct;62(10):e548-e558 [FREE Full text] [doi: 10.1097/JOM.0000000000001980] [Medline: 32769780]

27. Jimenez D, Nair S, Fu S, Schulz R, Czaja S. Reducing caregiver burden and stress in the workplace: a pilot intervention for working caregivers. The Gerontologist 2015;55:728-729. [doi: 10.1093/geront/gnv364.05]
28. Turgeman-Lupo K, Toker S, Ben-Avi N, Shenhar-Tsarfaty S. The depressive price of being a sandwich-generation caregiver: can organizations and managers help? European Journal of Work and Organizational Psychology 2020 May 18;29(6):862-879. [doi: 10.1080/1359432x.2020.1762574]
29. Lorenz F, Whittaker L, Tazzeo J, Williams A. Availability of caregiver-friendly workplace policies: an international scoping review follow-up study. IJWHM 2021 Jun 25;14(4):459-476. [doi: 10.1108/ijwhm-10-2019-0136]
30. Mofidi A, Tompa E, Williams A, Yazdani A, Lero D, Mortazavi SB. Impact of a caregiver-friendly workplace policies intervention: a prospective economic evaluation. J Occup Environ Med 2019 Jun;61(6):461-468. [doi: 10.1097/JOM.0000000000001564] [Medline: 31166280]
31. Dich N, Lange T, Head J, Rod NH. Work stress, caregiving, and allostatic load: prospective results from the Whitehall II cohort study. Psychosom Med 2015 Jun;77(5):539-547 [FREE Full text] [doi: 10.1097/PSY.0000000000000191] [Medline: 25984826]
32. Policy support to carers in selected places. Research Office-Legislative Council Secretariat. 2020. URL: https://www.legco.gov.hk/research-publications/english/1920rt07-policy-support-to-carers-in-selected-places-20200309-e.pdf [accessed 2025-03-21]
33. Ireson R, Sethi B, Williams A. Availability of caregiver-friendly workplace policies (CFWPs): an international scoping review. Health Soc Care Community 2018 Jan;26(1):e1-e14 [FREE Full text] [doi: 10.1111/hsc.12347] [Medline: 27071368]
34. Matos K, Galinsky E, Bond JT. 2016 National Study of Employers. SHRM. 2017. URL: https://cdn.sanity.io/files/ow8usu72/production/d73a7246cc3a3fef4ad2ece1e3d5aa4eaec2f263.pdf [accessed 2025-03-21]
35. Lee M, French C, Bai X, Yeung C, Taddese H. A Case Study of Caregiver-Friendly Workplace Policy in Hong Kong. London, UK: Imperial College London; 2018.
36. Sethi B, Williams A, Ireson R. Supporting caregiver employees: managers' perspective in Canada. International Journal of Workplace Health Management 2017 Feb 06;10(1):25-41. [doi: 10.1108/ijwhm-03-2016-0016]
37. Wong O, Chau B. The evolving role of filial piety in eldercare in Hong Kong. Asian J Soc Sci 2006;34(4):600-617. [doi: 10.1163/156853106778917790]
38. Leung VWY, Lam CM, Liang Y. Parents' expectations of familial elder care under the neoliberal Hong Kong society. Journal of Family Issues 2019 Sep 27;41(4):437-459. [doi: 10.1177/0192513x19879189]
39. 2016 Population By Census: Summary Results. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2017. URL: https://www.censtatd.gov.hk/en/data/stat_report/product/B1120094/att/B11200942016XXXXB0100.pdf [accessed 2025-03-21]
40. Welfare Services: Statistical Highlights - Residential care services for the elderly. Research Office-Legislative Council Secretariat. 2017. URL: https://www.legco.gov.hk/research-publications/english/1617issh22-residential-care-services-for-the-elderly-20170314-e.pdf [accessed 2025-03-21]
41. Panel on Education Subcommittee on Integrated Education. Legislative Council of the Hong Kong Special Administrative Region. 2014. URL: https://www.legco.gov.hk/yr13-14/english/panels/ed/ed_ie/reports/ed_iecb4-1087-1-e.pdf [accessed 2025-03-21]
42. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues in Clinical Neuroscience 2022 Apr 01;11(2):217-228. [doi: 10.31887/dcns.2009.11.2/hbrodaty]
43. Labour Force, Employment and Unemployment. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.censtatd.gov.hk/en/scode200.html [accessed 2025-03-21]
44. Wong SY, Zou D, Chung RY, Sit RW, Zhang D, Chan D, et al. Regular source of care for the elderly: a cross-national comparative study of Hong Kong with 11 developed countries. J Am Med Dir Assoc 2017 Sep 01;18(9):807.e1-807.e8. [doi: 10.1016/j.jamda.2017.05.009] [Medline: 28647578]
45. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
46. Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory (3rd ed). Thousand Oaks, CA: Sage Publications, Inc; 2008.
47. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: SAGE Publications, Inc; 2022.
48. Ritchie J, Spencer L. Qualitative Data Analysis for Applied Policy Research. In: Huberman AM, Miles MB, editors. The Qualitative Researcher's Companion. Thousand Oaks, CA: SAGE Publications, Inc; 2002:305-329.
49. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. JOAAG 2008;4(2):72-79 [FREE Full text]
50. Carer's Passport. Cabinet Office, Government People Group. URL: https://www.gov.uk/government/publications/carers-passport [accessed 2025-03-21]

51. Bondar J, Babich Morrow C, Gueorguieva R, Brown M, Hawrilenko M, Krystal JH, et al. Clinical and financial outcomes associated with a workplace mental health program before and during the COVID-19 pandemic. JAMA Netw Open 2022 Jun 01;5(6):e2216349 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.16349] [Medline: 35679044]
52. Peters P. Case study: awareness raising – Novel Zorg and Wonen, the Netherlands. Eurofound. 2011. URL: https://www.eurofound.europa.eu/en/case-studies-0 [accessed 2025-03-21]
53. Humer Z. Case study: awareness raising – Mercator, Slovenia. Eurofound. 2011. URL: https://www.eurofound.europa.eu/en/case-studies-0 [accessed 2025-03-21]
54. Barroso M. Case study: Work adjustments - Axa, Portugal. Ciência_Iscte. 2011. URL: https://ciencia.iscte-iul.pt/publications/case-study-work-adjustments---axa-portugal/48874 [accessed 2025-03-21]
55. Consultancy Study on Needs and Support Required of Carers of Elderly Persons and of Persons with Disabilities in Hong Kong. The Hong Kong Polytechnic University. 2022. URL: https://research.polyu.edu.hk/en/activities/consultancy-study-on-needs-and-support-required-of-carers-of-elde-2 [accessed 2025-03-21]
56. Kotter JP. Leading Change: Why Transformation Efforts Fail. Harvard Business Review. 1995. URL: https://heeoe.hee.nhs.uk/sites/default/files/leading_change_why_transformation_efforts_fail.pdf [accessed 2025-03-21]
57. Huque AS, Jongruck P. Civil service reforms in Hong Kong and Thailand: similar goals, different paths. Public Administration and Policy 2020 Aug 13;23(2):111-123. [doi: 10.1108/pap-03-2020-0015]
58. Li E. A case study of the critical success factors for organizational change of a public listed corporation. Public Administration and Policy 2018 Nov 07;21(2):152-165. [doi: 10.1108/pap-10-2018-010]
59. Wong M, Yeung D. Permission-based leadership and change management in Hong Kong's nongovernment organisations. Nang Yan Business Journal 2014;3(1):67-91 [FREE Full text] [doi: 10.1515/nybj-2015-0006]
60. Nesbit PL, Lam E. Cultural adaptability and organizational change: a case study of a social service organization in Hong Kong. Contemporary Management Research 2014;10(4):12186. [doi: 10.7903/cmr.12186]
61. Support to Small and Medium Enterprises. Trade and Industry Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.tid.gov.hk/english/smes_industry/smes/smes_content.html [accessed 2025-03-21]
62. Scheme on Living Allowance for Carers of Elderly Persons from Low-income Families. Social Welfare Department-The Government of Hong Kong Special Administrative Region. 2022. URL: https://www.swd.gov.hk/en/pubsvc/elderly/cat_careersupp/allowanceforcarersoftheelderly/ [accessed 2025-03-21]
63. Hong Kong as a Knowledge-based Economy - A Statistical Perspective. Census and Statistics Department-The Government of Hong Kong Special Administrative Region. 2019. URL: https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=B1110009&scode=590 [accessed 2025-03-21]
64. Fact Sheet: Voluntary Leave Transfer Program. US Office of Personnel Management. URL: https://www.opm.gov/policy-data-oversight/pay-leave/leave-administration/fact-sheets/voluntary-leave-transfer-program/ [accessed 2025-03-21]
65. Bradbury MD. Voluntary shared leave: an exploratory analysis. Public Personnel Management 2019 Nov 27;49(4):532-546. [doi: 10.1177/0091026019889378]
66. Carers Canada. URL: https://www.carerscanada.ca/ [accessed 2025-03-21]
67. Key facts and figures about caring. Carers UK. 2022. URL: https://www.carersuk.org/news-and-campaigns/press-releases/facts-and-figures [accessed 2025-03-21]
68. About Taiwan Association of Family Caregivers. Taiwan Association of Family Caregivers. URL: https://www.familycare.org.tw/about [accessed 2025-03-21]
69. Compassionate care, family caregiving and bereavement leaves. University of British Columbia. URL: https://hr.ubc.ca/benefits/benefit-plan-details/vacation-leaves/compassionate-care-compassionate-or-bereavement-leave [accessed 2025-03-21]
70. Tucker EM, Vosko LF. Designing Paid and Protected Employment Leaves for Short-Term Sickness and Caregiving. IRPP Insight. 2021. URL: https://irpp.org/research-studies/designing-paid-and-protected-employment-leaves-for-short-term-sickness-and-caregiving/ [accessed 2025-03-21]
71. Act on Childcare Leave, Caregiver Leave, and Other Measures for the Welfare of Workers Caring for Children or Other Family Members. Japanese Law Translation. URL: https://www.japaneselawtranslation.go.jp/en/laws/view/2438/en [accessed 2025-03-21]
72. Yuhara E. Japan's 'Nursing Murders.'. Nippon. 2017. URL: https://www.nippon.com/hk/currents/d00251/?pnum=2# [accessed 2025-03-21]
73. Parker C, Scott S, Geddes A. Snowball Sampling. Thousand Oaks, CA: SAGE Publications Ltd; 2019.
74. Sargeant J. Qualitative research part II: participants, analysis, and quality assurance. J Grad Med Educ 2012 Mar;4(1):1-3 [FREE Full text] [doi: 10.4300/JGME-D-11-00307.1] [Medline: 23451297]
75. Roberts JK, Pavlakis AE, Richards MP. It's more complicated than it seems: virtual qualitative research in the COVID-19 era. International Journal of Qualitative Methods 2021 Mar 25;20:16094069211002959. [doi: 10.1177/16094069211002959]

Abbreviations

CE: carer-employee

CFWP: caregiver-friendly workplace policy
ESG: environmental, social, and governance
MPF: mandatory provident fund
NGO: nongovernmental organization

Edited by P Jansons; submitted 28.10.24; peer-reviewed by E Okoyeocha, Y Jiang; comments to author 15.01.25; revised version received 13.02.25; accepted 11.03.25; published 31.03.25.

Please cite as:
Lee MMS, Yeoh EK, Wong ELY
Employers' Perspectives of Caregiver-Friendly Workplace Policies for Caregiver-Employees Caring for Older Adults in Hong Kong: Thematic Analysis
JMIR Aging 2025;8:e68061
URL: https://aging.jmir.org/2025/1/e68061
doi:10.2196/68061
PMID:

Original Paper

Creating User Personas to Represent the Needs of Dementia Caregivers Who Support Medication Management at Home: Persona Development and Qualitative Study

Anna Jolliff[1], MS, LMHC; Priya Loganathar[2], PhD; Richard J Holden[3], PhD; Anna Linden[2], PhD; Himalaya Patel[3,4], PhD; Jessica R Lee[5], MS; Aaron Ganci[6], MFA, PhD; Noll Campbell[7,8,9], MSc, PharmD; Malaz Boustani[8,9,10], MPH, MD; Nicole E Werner[1], PhD

[1]Center for Research and Innovation in Systems Safety, Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN, United States

[2]Department of Industrial & Systems Engineering, University of Wisconsin - Madison, Madison, WI, United States

[3]Department of Health & Wellness Design, School of Public Health, Indiana University Bloomington, Bloomington, IN, United States

[4]Health Systems Research Center for Health Information and Communication, Richard L Roudebush Veterans Affairs Medical Center, United States Department of Veterans Affairs, Indianapolis, IN, United States

[5]Department of Human Systems Engineering, Ira A. Fulton Schools of Engineering, Arizona State University, Mesa, AZ, United States

[6]Herron School of Art + Design, Indiana University Indianapolis, Indiana University Indianapolis, Indianapolis, IN, United States

[7]Department of Pharmacy Practice, Purdue University, West Lafayette, IN, United States

[8]Indiana University Center for Aging Research, Regenstrief Institute, Indianapolis, IN, United States

[9]Center for Health Innovation and Implementation Science, Indiana University, Indianapolis, IN, United States

[10]Department of Medicine, Indiana University, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Anna Jolliff, MS, LMHC
Center for Research and Innovation in Systems Safety
Department of Anesthesiology
Vanderbilt University Medical Center
2525 West End Avenue
Suite 800
Nashville, TN, 37203
United States
Phone: 1 615 936 6598
Email: anna.jolliff@vumc.org

Abstract

Background: Caregiver-assisted medication management plays a critical role in promoting medication adherence and quality of life for people living with Alzheimer disease or related dementias (ADRD). The current landscape of digital and nondigital interventions to support medication management does not meet caregivers' needs, contexts, and levels of technological proficiency. Intervention development can be facilitated using personas or data-driven archetypes that represent end users' traits relevant to solution design.

Objective: This study aims to understand the strategies and unmet needs of ADRD caregivers who manage medications and use this understanding to create personas that can inform customized caregiver interventions.

Methods: Participants were self-identified primary caregivers of people with ADRD living with or near the care recipient. Virtual contextual inquiry was completed in three stages: (1) enrollment interview, (2) virtual observation over a 1-week period, and (3) postobservation interview. Codebook thematic analysis of interview transcripts was used to identify dimensions of caregivers' approaches to medication management. A reflexive, team-based affinity diagramming approach was used to identify attributes within these dimensions and group attributes into personas.

Results: Participants (N=25) were aged 62.32 (SD 11.86) years on average, and 17 (68%) of them were female. Caregivers varied across 6 dimensions relevant to medication management: strategies for medication acquisition, medication storage and organization, medication administration, monitoring the care recipient for symptoms, communication with care network regarding

medication, and acquiring information about medication. Three personas were created to represent the observed strategies, unmet needs, and levels of technology use related to medication management: Checklist Cheryl, in Control; Social Sam, Keeps it Simple; and Responsive Rhonda, Stays Relaxed.

Conclusions: Caregivers in this study demonstrated a range of characteristics and values that informed their approach to medication management. They used a combination of technology-based strategies and strategies situated in their physical environments to manage medications. The personas created can be used to inform interventions, such as digital tools, that address caregivers' unmet needs.

(JMIR Aging 2025;8:e63944) doi:10.2196/63944

KEYWORDS

medication adherence; user-centered design; digital health technology; informal caregivers; dementia; artificial intelligence; AI

Introduction

Overview

Every year, family and friends of people living with Alzheimer disease or related dementias (ADRD) around the globe provide 82 billion hours of care [1]. These family caregivers assist in promoting the health and safety of the persons living with ADRD, including helping them manage complex medication routines [2-5]. A US-based survey of family caregivers found that medication management is a challenging process due to the considerable time involved, the complexity of routines, fear of making a mistake, and care recipients' resistance to taking medication [6]. Caregivers of people living with ADRD may face unique challenges; for example, an Australian study of family caregivers found that people living with ADRD may not recognize their need for assistance with managing medication and may feel embarrassed to accept help [7]. International reviews of the literature further suggest that caregivers of people living with ADRD lack training and resources to support the breadth of activities associated with medication management [5,8]. Thus, ADRD caregivers would benefit from interventions that are tailored to their needs, contexts, and levels of technological proficiency to facilitate safe and successful medication management [9]. Intervention development can be facilitated using personas or data-driven archetypes that represent end users' traits relevant to solution design.

Background

Challenges With Caregiver-Assisted Medication Management

Medication management is a complex, multifaceted process [10]. One aspect of caregiver-assisted medication management (henceforth, medication management) is maintaining a constant supply of medications, which is made more complicated by differences in prescription pack sizes, delays in prescription refill, receiving incomplete prescriptions, and return visits to the pharmacy when concerns or errors are detected [10]. Administering medications is complicated by the intricacies of many medication routines, frequent changes to these routines (including the dose, number, and form of medications), and both caregivers' and care recipients' forgetfulness to take medications [5,10]. Caregivers are also responsible for making judgments, such as when to withhold, increase, decrease, or discontinue medication, which involves monitoring the care recipient's symptoms, medication side effects, and whether medications are having their intended effect [10]. Finally, communication with the care team is a key but ambiguous component of medication management; caregivers may struggle to know when to reach out to prescribers or to feel comfortable reaching out and may receive incomplete information about medications from health care providers (eg, side effects to look out for or changes to a prescription) [10,11]. Thus, medication management can represent significant and taxing work for family caregivers.

Caregiver-Assisted Medication Management and Well-Being

Medication management has implications for the health and well-being of both the care recipient and caregiver. For the care recipient, international research spanning the United States, Europe, Canada, Japan, and Australia has linked suboptimal medication management with poorer health [12], poorer management of comorbid conditions [13], costly hospitalizations [14], and premature institutional placement [15]. With assistance from caregivers, medication adherence rates among those with ADRD can match those of their peers without cognitive impairment [3]. Among caregivers, performing medical tasks has been associated with positive outcomes, such as gaining new skills and feeling closer to and less worried about the care recipient [6]. However, medication management has also been associated with greater caregiver burden and stress in international research [7,16-18]. In the United States, nearly 20% of family caregivers who perform medical tasks for the care recipient describe fear of making a mistake, 32% report continuously monitoring for something to go wrong, and 40% report feeling "down, depressed, or hopeless" as an outcome of performing these tasks [6]. Therefore, it is critical to understand the unmet needs of caregivers managing medications as the first step in maximizing the positive and mitigating the negative outcomes for care recipients and caregivers alike.

Limitations of Existing Medication Management Tools

Many tools already exist to support medication management, from low-tech tools (eg, plastic cups and sandwich bags) to high-tech tools (eg, electronic pillboxes) [9]. Some caregivers' needs are met by their existing approaches, and, for them, no intervention may be needed [9]. However, other caregivers report that current low-tech tools, such as pillboxes and reminder notes, do not meet their needs [19]. Pillboxes may not accommodate more complex medication routines, and reminder notes may be ignored over time [19]. While digital interventions

abound, a recent systematic review of mobile apps available in Google Play or iTunes intended to assist with medication management suggested that these, too, may not meet users' needs [20]. The reviewed apps were largely low quality and missing desirable features, such as refill reminders and information about adherence [20]. In a US-based study of family caregivers, participants perceived existing digital interventions for medication management as too elaborate or overwhelming [9]. Thus, it is evident that current interventions to support medication management do not meet the full range of needs, contexts, and interests in technology of the current population of ADRD caregivers. By examining the strategies currently used by caregivers to manage medication, we can better understand the features of interventions that would support medication management [21].

A Call for Customized Solutions

Each caregiver's medication management needs and contexts are different [22]. Needs may vary based on the nature of the routine [23-25], qualities of the care recipient [6,8,26,27], and caregiver characteristics [28,29]. The routine itself is shaped by the variety of medications prescribed, the number of prescribers, and the use of multiple pharmacies [23-25]. Qualities of the care recipient include the stage and symptoms of dementia [8] and any co-occurring medical conditions [6], both of which may change over time, in turn, prompting changes to their medication routine [26,27]. Caregivers vary in their health literacy, access to information, commitment to self-care, and valuing of the care recipient's independence [28], all of which may affect how they manage medication. Caregivers also differ on the degree to which they receive support from others with caregiving, including medication management [8]. One analysis that identified typologies of caregiver support networks found that male caregivers and low-income caregivers in the United States were more likely to be in low support caregiving networks, that is, networks in which caregivers receive less support for themselves (eg, emotional support and support with housework) or for the care recipient (eg, personal care and transportation) [29]. In summary, the intricacies of medication management intersect with the characteristics of caregivers and care recipients to yield distinct requirements of supportive interventions.

Translating Unique Needs Into Solution Design

There is an outstanding need to understand the variable needs of caregivers managing medication on behalf of people living with ADRD and design solutions to meet those needs. Furthermore, in the design of solutions to address these caregiving challenges, it is critical to involve caregivers themselves [30]. User-centered design (UCD) is an internationally recognized approach for integrating future end users into the design of interventions that aim to meet their unique and complex needs [31]. UCD is defined as "an approach to interactive systems development that aims to make systems usable and useful by focusing on the users, their needs and requirements" [31]. UCD engages representatives of the user population in the design of interventions, resulting in interventions that are usable, useful, relevant, and adaptable to caregivers' needs [21,32]. UCD has proven valuable to the design of solutions in the health care domain across the globe to meet diverse patient population needs [30,33-35].

Within the field of UCD, persona design is a specific method in which real participant data are analyzed and distilled into archetypal end-user representations or personas [36]. Personas can be used to represent the range of key user dimensions regarding a certain process (eg, medication management), including relevant demographic characteristics, strategies they use to do their work, and any unmet needs they encounter [37]. In the context of intervention development, personas have certain advantages over other products of qualitative inquiry, such as theory, themes, or concept maps; they offer concise, data-driven representations of end users, allowing designers to ask and answer the question, *What do our end users need?* [36]. Personas may be particularly helpful for informing digital caregiving designs, including home monitoring systems, automated support with caregiving tasks, social robots that provide comfort to care recipients, and social networks that provide support to caregivers [38]. Persona development is a life cycle; initial personas are somewhat crudely defined by their distinguishing features (conception), enriched with storytelling and stakeholder feedback (gestation), and iteratively refined as they are applied (adulthood) [36]. Previous studies in the United States [39,40], Singapore [41], and Germany [42] have used personas to depict important variations among the target populations of interventions and to infer the necessary intervention features to meet users' unique needs.

Study Purpose

This study aimed to answer the following research questions (RQs):

- RQ1—What are the strategies used by caregivers of people living with ADRD who manage medication?
- RQ2—What are the unmet needs of caregivers of people living with ADRD related to medication management?
- RQ3—How can the answers to RQs 1 and 2 be represented as personas to inform intervention development?

Methods

Ethical Considerations

This study was subject to an expedited ethics review and received approval from the University of Wisconsin–Madison Institutional Review Board (2021-0339). Participants provided verbal consent to participate and were told that they could withdraw consent at any time. All data were deidentified before analysis. Participants received US $25 for completing the enrollment interview and another US $25 for completing the postobservation interview.

Study Design

To identify the approaches and unmet needs related to the daily work of medication management for caregivers of people living with ADRD, we adapted a design research method, contextual inquiry, to a virtual form, which we have termed virtual contextual inquiry (VCI) [43,44]. Traditional contextual inquiry combines principles of ethnography with cognitive task analysis, seeking to understand how tasks are completed in naturalistic

environments, including the implicit and tacit knowledge obtained from repeatedly performing contextualized work [45]. In traditional contextual inquiry, researchers observe people in their natural environments while they perform tasks of interest (eg, medication management) as they normally would. Interview probes are inserted unobtrusively before, during, and after observation. In VCI, participants send the research team multimedia messages (eg, photos and textual descriptions of work) during or soon after completing the work [43]. These multimedia messages allow for nearly real-time observation of tasks and serve as a cognitive aid during the postobservation interview [46,47]. As with in-person contextual inquiry, in VCI, data collection is structured across 3 phases: before, during, and after observation.

In this study, VCI was selected over traditional contextual inquiry for myriad reasons. First, medication management takes place in various physical spaces and may vary depending on the day [10]. In order to gain a more accurate picture of the work, observation must be close and sustained. However, to observe participants in person over multiple days would be costly and time consuming for the researcher as well as inconvenient and even uncomfortable for participants [48]. Virtual observation thus strikes a balance between practicality and comprehensiveness [46,49]. Virtual observation may also improve the accuracy of data; participants who are less conscious of being observed may be less likely to alter their typical behaviors [50,51]. Furthermore, virtual observation inherently creates artifacts that can be repeatedly referenced throughout analysis, meaning researchers are less dependent upon memory or written notes. Finally, VCI allows for widespread recruitment, rather than recruiting solely from areas within geographic proximity to the research institution, and permits scholarship to continue in the face of circumstances that prevent in-person observation (eg, the COVID-19 pandemic) [48].

Conceptual Framework

Our conceptual framework for medication management, which influenced the enrollment and postobservation interview process, was based on the findings of Smith et al [10] related to contexts in which medication-related problems can occur for family caregivers managing medications. We were also guided by previous research on crafting personas for intervention design and thus inquired about participants' unmet needs and technology use [36,52].

Study Sample

This study used 2 different methods to enroll a convenience sample of caregivers of people living with ADRD. In the first method, participants were recruited from the Wisconsin Alzheimer's Disease Research Center's registry of self-identified primary caregivers and in the second method from the research team's ADRD caregiver registry. The eligibility criteria included (1) self-identifying as a primary caregiver for a person living with ADRD; (2) aged ≥18 years; (3) managing medication for a person living with ADRD, as defined by our conceptual framework; (4) currently living with or close to a person living with ADRD; (5) having access to a cellphone during the study period; and (6) speaking English.

Study Procedure

Recruitment

Participants who were recruited from the Wisconsin Alzheimer's Disease Research Center emailed the study team if they were interested in participating. The study information sheet was then shared with the participant, and a phone screening was scheduled. Members of the study team's caregiver registry were first contacted by phone, as they had previously indicated a willingness to be contacted regarding research opportunities. For both recruitment methods, after the screening call was complete and participants provided verbal consent to participate, the enrollment interview was scheduled. Those eligible and willing to participate completed the VCI process, which consisted of three stages: (1) enrollment interview, (2) virtual observation over a 1-week period, and (3) postobservation interview. The interviewer (PL) and participants were unknown to each other before recruitment, and participants knew only the interviewer's professional status as a doctoral student.

Enrollment Interview

The enrollment interview was a 1-hour audio-recorded session that took place over a secure videoconferencing software. During the interview, participants were introduced to the text messaging procedure to be performed during the observation period. Participants were provided with examples of multimedia text messages that would meet study criteria, invited to add the study team's phone number as a contact in their phone, and practiced submitting a message to the study phone. Participants were also provided with an instructional video that detailed how to submit a message. The interviewer (PL) cautioned participants against submitting messages that contained identifying information. Following this, participants indicated 2 times in the morning and evening at which they would like to receive prompts to submit a message from the study team. The interviewer scheduled a final interview that would occur after the observation period and approximately 8 to 10 days from the enrollment interview. Finally, participants completed a demographic survey and were asked to describe their daily, weekly, and monthly medication management activities. On the basis of this information, the interviewer created an initial list of potential persona dimensions (eg, *medication administration*), which were brought to the research team for discussion and refinement.

Observation Period

During the observation period, participants sent multimedia text messages (eg, textual descriptions, still photos, videos, and audio files) that described their medication management to a secure phone dedicated exclusively to study-related communication. Participants were asked to send at least 2 multimedia text messages daily for 7 days, capturing the medication management activities they performed. At the beginning and end of each day, participants received a text message reminding them to send messages. All messages were deidentified to protect the identity of the participants and uploaded to a shared and secure cloud-based drive. After each participant's observation period, the research team met to review their observation data and made notes of any multimedia

messages that needed more clarification in the postobservation interview (eg, if the tasks pictured were not clear or if any essential context was missing). In this meeting, the research team also used the caregivers' messages to summarize their approach to medication management, which would later be presented to the caregiver for feedback in the postobservation interview.

Postobservation Interview

We conducted a 1-hour postobservation interview via videoconferencing software within 8 to 10 days after the enrollment visit. During the interview, an interviewer trained in contextual inquiry reviewed the participant's multimedia messages and probed for detail. The interview followed the standard procedure for contextual inquiry in that questions did not follow a predetermined interview guide but instead were specific to the data supplied by participants [53]. For example, one participant (P15) shared an image of the spreadsheet they used to track medications. The interviewer asked, "Can you talk a little bit about how you started using that [spreadsheet] or how you came up with that?" Another caregiver (P33) shared over text that the care recipient had missed a dose of their medication, which meant that the caregiver had to remind them to take it. The interviewer asked, "How did you know he missed his medication, and then how did you go about reminding him?" Finally, the interviewer presented the participant with their summarized approach to medication management for reflection and refinement. The interview was audio recorded and transcribed for analysis.

Analysis

Overview

There is no single, standardized approach to persona development [39-42,53,54]. Our approach combined positivist and nonpositivist approaches [55]. We first used a structured, positivist approach to identify dimensions of medication management. We used this approach under the assumption that there were a finite number of medication management activities in which caregivers could engage, many of which could have been anticipated before analyses. It was reasonable to assume that 2 coders could identify the same activities, regardless of their perspectives.

With that in mind, we first used Microsoft Excel to conduct a codebook thematic analysis of postobservation interview transcripts to identify dimensions of medication management [55]. Initial dimensions were those generated by the interviewer in the enrollment interview. Four coders (PL, AL, JRL, and Mengwei Tang) coded the first 2 postobservation interview transcripts as a team. During that process, the coders refined the initial codebook, adding or modifying dimensions, definitions, and examples. The initial codebook was then brought to the entire research team and refined through consensus-based discussion. The codebook was then applied separately to the remaining transcripts by the 4 coders, such that each transcript was coded by 2 coders. To promote reliability, coders met weekly to discuss their coding and resolve any discrepancies through consensus-based discussion [56]. The coding team also met regularly with the full research team to discuss coding and resolve any discrepancies [56]. Any emergent codes or proposed changes to the codebook were brought to the entire research team for discussion and approval. The resulting final codebook represented an enumeration of the initial persona dimensions.

We then adopted a nonpositivist approach to move from dimensions to interpretive stories, in this case, toward personas and their attributes. This shift was consistent with a philosophical assumption held by the research team that, while it is possible to identify objective and discrete dimensions of medication management (eg, *medication acquisition*), we were unlikely to identify discrete types of people through our analysis. The enumeration of attributes within these dimensions and the assignment of attributes to personas were assumed to be subjective and interpretive, a product of the minds that performed the work [55].

To move from dimensions to attributes, we identified passages coded to a specific dimension and created a sentence-length paraphrase of each passage. These paraphrases were transferred to virtual sticky notes, color-coded by participant ID, and grouped by perceived similarity in an iterative team-based affinity diagramming process. For example, within the dimension *approach to medication acquisition*, sticky notes perceived as describing a proactive approach to medication acquisition were grouped together to form an attribute. If it was perceived that participants with this proactive approach to medication acquisition tended to use a strategic approach to storing medications, these attributes were spatially positioned near each other on the virtual sticky board. Each participant was assigned a distinct color sticky note. If, for example, three distinct colors appeared in the same regions of the whiteboard, it practically meant that those three participants shared an attribute. When multiple participants seemed to share multiple attributes, it warranted the creation of a persona. A process of crystallization was used to add depth and nuance to the final personas [57]. Transcripts were critically reviewed by a research team member (PL) to look for information that could add detail to the developed personas, ensure all perceived attributes had been represented by a persona, and check that no persona too closely resembled any one participant. The cowriting of this manuscript allowed the multidisciplinary research team to compare understandings of how attributes fit or did not fit within each persona. Member reflections (a Big Q alternative to member checking) were also used to refine personas [58]. The research team shared drafts of personas with 7 ADRD caregivers who were not involved in the study and augmented the personas based on their feedback. This approach was not intended to *verify* personas but rather add further depth and nuance to the research team's interpretations.

Research Team Positioning

Members of the research team held professional positions that inevitably influenced their perceptions of attributes and the creation of personas. Four members of the research team were faculty (RJH and NEW) or doctoral students (PL and AL) in the field of human factors engineering, which meant that attributes likely to shape or pose a barrier to work performance were naturally emphasized, along with attributes with

implications for product design. The lead author of this paper (AJ) was a licensed mental health counselor with training in theories of personality and thus brought assumptions about the likelihood that certain attributes would or would not co-occur. The collaboration of a pharmacist (NC), geriatrician (MB), and a research health scientist specializing in human-computer interaction (HP) added depth of understanding to topics relevant to the RQ.

Results

Overview

We reached out to 40 caregivers, and 25 (62%) of them enrolled in the study.

Caregivers who chose not to enroll largely did not respond to our outreach; of those who did respond, 10% (4/40) were ineligible, 3% (1/40) were unable to find time to schedule the enrollment interview, and 3% (1/40) were not interested in sending multimedia messages.

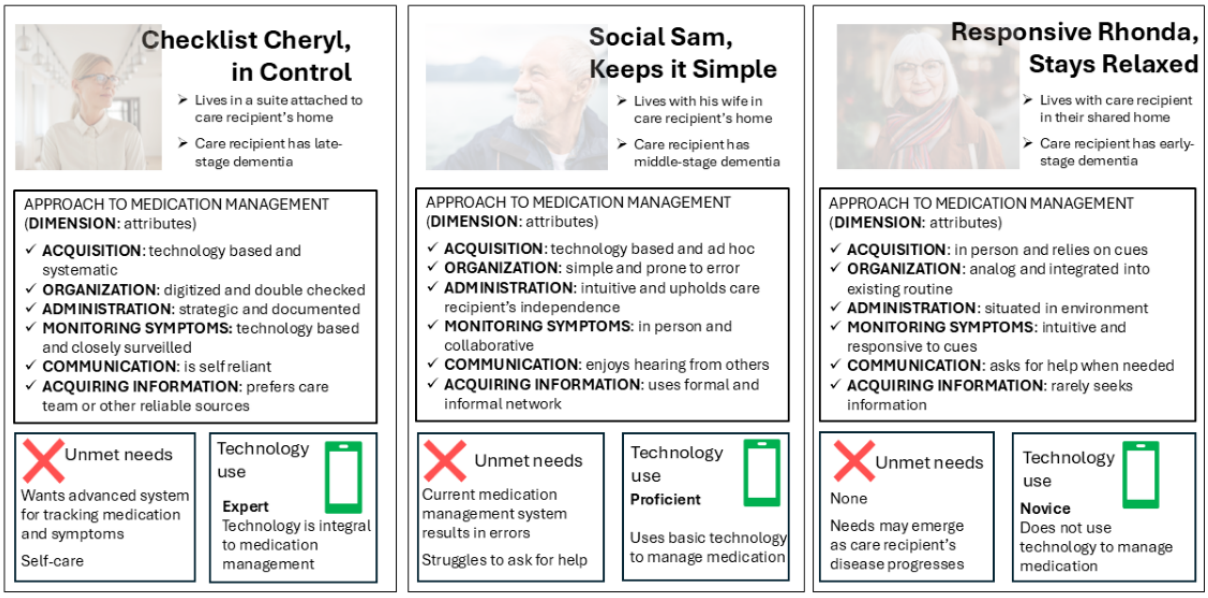
On average, participants were aged 62.3 (SD 11.9) years; 68% (17/25) of the participants were female, and all (25/25, 100%) identified as White and not of Hispanic or Latino origin. The care recipients, on average, were aged 76.7 (SD 11.1) years, and 12% (3/25) were aged <65 years. Among care recipients, 48% (12/25) were female, 92% (23/25) were White, 4% (1/25) were Black or African American, and 4% (1/25) were of Hispanic or Latino origin. All (25/25, 100%) participants demonstrated multiple medication management activities over the course of the 7-day observation period, and 96% (24/25) of the participants submitted messages on at least 5 of the 7 days. Participant characteristics are summarized in Table 1.

Table 1. Participant demographics (N=25).

Characteristics	Values
Sex, n (%)	
Female	17 (68)
Male	8 (32)
Age (y), mean (SD)	62.3 (11.9)
Race, n (%)	
Black or African American	0 (0)
White	25 (100)
Ethnicity, n (%)	
Hispanic or Latino	0 (0)
Non-Hispanic or Latino	25 (100)
Highest educational attainment, n (%)	
High school diploma or equivalent	1 (4)
Technical school, vocational training, or community college	5 (20)
4-year college	8 (32)
Postcollege education	11 (44)
Employment status, n (%)	
Full time	4 (16)
Part time	3 (12)
Not working	4 (16)
Retired	14 (56)
Annual income (US $), n (%)	
<20,000	1 (4)
20,000-39,999	3 (12)
40,000-$59,999	5 (20)
60,000-79,999	6 (24)
80,000-99,999	5 (20)
>100,000	2 (8)
Undisclosed	3 (12)
Length of time managing medications for care recipient (y), mean (SD)	5.0 (3.1)
State of residence, n (%)	
Wisconsin	16 (64)
California	5 (20)
Other	3 (12)
Undisclosed	1 (4)
Care partner's relationship to the care recipient, n (%)	
Adult child or child-in-law	8 (32)
Spouse or significant other	16 (64)
Other (mother and father)	1 (4)
Care recipient's living arrangement, n (%)	
Lives in the primary caregiver's home	20 (80)
Lives in the care recipient's own home	4 (16)
Other (independent living facility)	1 (4)

We derived 3 personas: Checklist Cheryl, in Control; Social Sam, Keeps it Simple; and Responsive Rhonda, Stays Relaxed. These personas varied across 6 dimensions relevant to medication management: approach to medication acquisition and organization, medication administration, monitoring the care recipient for symptoms, communication with the care network regarding medication, and acquiring information about medication. The presentation of the personas given subsequently is organized by their approach to each dimension of medication management, with illustrative examples and quotations from real study participants. Persona profiles are presented in Figure 1.

Figure 1. Three persona profiles representing caregivers' medication management for patients with Alzheimer disease or related dementias.



Checklist Cheryl, in Control

Overview

Checklist Cheryl, in Control, is characterized by her desire to do things on her own, be in control, and create reliable and precise systems and regimens. Cheryl lives in a suite that is separate from but connected to the home of the care recipient, who has late-stage dementia.

Approach to Acquiring and Organizing Medication

Checklist Cheryl has an organized, technology-based system for acquiring and maintaining a constant supply of medications. She sets up automated refills and mail-based delivery for all medications. She organizes medication in an automated medication dispenser, which she fills every 2 weeks. When filling the dispenser, she verifies the dosage in a notebook that describes each medication as well as a large, printed list of the medications with pictures and a description of each that she keeps on the wall for her reference.

Approach to Administering Medication

Checklist Cheryl continuously improved upon her strategies for medication administration and documents the effectiveness of these strategies in a journal. One participant described iterative strategizing to encourage the care recipient to take her medication:

> *[The care recipient] didn't like [taking medication], and she made bad faces. So, I thought, one day I'll just give her something extra. It started, I mean, we've done like Pringles and different types of chips and crackers and all that kind of stuff.* [P1]

Approach to Monitoring for Symptoms

Checklist Cheryl closely monitors the care recipient for symptoms and writes down her observations in a journal to track changes in behavior. She installs a video surveillance system in the care recipient's home to better monitor any symptoms or changes in the care recipient's behavior:

> *If someone goes in the laundry room, it just brings up my cam and app, and then I can just click on the laundry room, and I can quickly look and see who went in there and what they're doing.* [P1]

Approach to Acquiring Medication Information

Checklist Cheryl seeks formal advice about medications. She contacts the care recipient's health care team with medication questions:

> *The medications he's getting now are from the neurologist. So if I see something is wrong, I call the neurologist for changes or advice.* [P7]

Alternatively, Cheryl may do her own research online. Cheryl does not rely on support groups, friends, or family for medication information.

Approach to Communicating With Care Network

Checklist Cheryl was unlikely to ask her formal care team or network of family and friends for help with medication management:

> *I do all of it. And I order [the medications] myself. Hospice said they would order [the care recipient's medications], but I order [the care recipient's medications] at the same time I order my husband's,*

> *so I might as well do it all. I'm kind of a control freak.* [P40]

Technology Use

Checklist Cheryl is a technology expert. Technology is integral to her medication refill system, medication dispenser, information-seeking approach, and the surveillance system she uses to monitor the care recipient's behavior.

Unmet Needs

Checklist Cheryl wants a better system for tracking the care recipient's changing symptoms and side effects over time. She is also unsure when to communicate symptoms or questions to the formal care team versus relying on the answers she can find through her own research. Finally, she finds that medication management is all consuming and leaves little time for self-care, which often leads to stress.

Social Sam, Keeps It Simple

Overview

Social Sam, Keeps it Simple, is characterized by his attributes of being collaborative, seeking efficiency, and valuing the care recipient's independence. Sam and his wife live in the home of the care recipient, who has middle-stage dementia.

Approach to Acquiring and Organizing Medication

Social Sam has a simple system for monitoring and restocking the medication supply. When he sees a pill bottle running low, he places it near his office computer:

> *[The medication bottles] will be right in front of me on top of the, my desk where my computer is...I'll just get on the website and put in the order right then and there.* [P8]

Sam refills an analog medication organizer on the same evening each week. When possible, Sam incorporates the storage of medication into the care recipient's existing routine. For example, one participant described storing inhalers in the care recipient's backpack:

> *When [the care recipient] used to work, he had that backpack with him 100% of the time, and so that's where the inhalers were. And that, I don't want to change that for him.* [P10]

Approach to Administering Medication

Social Sam's routine for administering medication is so consistent that it has become second nature:

> *It's the first thing that I do when I get up...the first thing I do is grab the pill. I also have medication I'm giving the cats, so I grab that at the same time and put it all out on the coffee table, where I keep it during the day.* [P8]

Because Sam values the care recipient's independence, he is subtle when checking whether medication has been taken:

> *I just pretend I'm doing something else and watch [him take the medication].* [P10]

Approach to Monitoring for Symptoms

Social Sam collaborates with his wife, who is also a member of his care network, to monitor the care recipient for symptoms or changes in behavior:

> *One of the things that's happened is [the care recipient] has gotten GERD, probably from all the medication that she's taking, upsetting her stomach. So she's belching more...And so [my wife] and I talk all the time...There's a constant that we check in, how is she? Is that medication bothering her?* [P25]

To monitor the care recipient's pain symptoms, Sam takes weekly glances at the as-needed pain medication bottle to see if there are fewer pills than there were the previous week.

Approach to Acquiring Medication Information and Communication With the Care Network

Social Sam enjoys receiving information and help from others, both clinicians and nonclinicians. For example, he prefers to pick up medications in person so he can ask the pharmacist questions. Sam uses informal sources of medication information, such as support groups and Facebook:

> *I try to draw on everyone and try to get the best information possible to make the right decision.* [P2]

Technology Use

Social Sam is technology proficient. He uses text messaging to communicate with his wife regarding medication management and uses phone reminders to prompt himself to do medication-related tasks.

Unmet Needs

Social Sam finds his current medication refill reminder system to be too simplistic, not accounting for medications with different refill dates. This insufficiently nuanced system results in Sam occasionally forgetting to refill medications and, consequently, missing doses. Though Sam is naturally collaborative, he finds coordinating medication management activities with other family members and formal care team members to be challenging, as he does not like to bother people.

Responsive Rhonda, Stays Relaxed

Overview

Responsive Rhonda, Stays Relaxed, is characterized by the attributes of being laid-back, nonintrusive, and responding to (rather than preparing for) challenges with medication management as they arise. Rhonda lives with the care recipient, who has early-stage dementia.

Approach to Acquiring Medication

Responsive Rhonda responds to overt reminders and alerts to order medication. For example, she may see the medication supply and notice it is low, or she may look into refilling medications when at the pharmacy for other reasons:

> *Like if I'm going in to pick something up in, at the pharmacy, I will just show them his number, his medical number also and just say, hey, is there anything that's available for pickup?* [P10]

Approach to Organizing and Administering Medication

Responsive Rhonda relies on nontechnological cues to remind herself to organize and administer medications. She stores pills where she regularly performs her nighttime routine:

> *The pill is just part of the setup that I do every night, so it's actually not too hard to remember to put that out.* [P9]

As part of existing daytime routine of Rhonda, she walks past the medication she set out and notices if the care recipient has taken the medications or if they need to be reminded.

Approach to Monitoring for Symptoms

Responsive Rhonda does not have a formal system to monitor the care recipient for symptoms, instead relying on intuition:

> *And if I feel like, this is more intuitive, is if I feel like he's [using Tylenol] much more than normal, then that will trigger me to go check on, you know, to ask, dig in a little more about his pain.* [P1]

Rhonda also relies on the care recipient to report symptoms to her.

Approach to Acquiring Medication Information and Communication With the Care Network

Responsive Rhonda is comfortable relying on others in the care network for help, although at this stage in the care recipient's disease, she rarely needs help. When Rhonda is away, she asks other family members to put the medications out for the care recipient and ensure that they take them. Rhonda seldom sees the need for medication information or advice, but when she does have questions, she is comfortable asking the care recipient's care team.

Technology Use

Responsive Rhonda is a technology novice. Although proficient with mobile phone calls, she does not think of incorporating technology into medication management.

Unmet Needs

Responsive Rhonda does not currently perceive any unmet needs related to medication management.

Discussion

Principal Findings

In this study, we leveraged the UCD method of VCI to understand ADRD caregivers' approaches to and needs for medication management. Caregivers exhibited a range of characteristics and values that informed their approach to medication management.

In addition, caregivers used a combination of technology-based strategies and strategies situated in their physical environments to manage medications. However, these personas must be understood with awareness of the relatively homogenous sample, which may limit generalizability. The average caregiver in this study was a retired or unemployed White woman in her 60s, with college or postcollege education, caring for someone in their 70s who lived in her home in the United States. These personas can be used to inform the UCD of interventions for caregivers with similar attributes by enabling designers to build empathy with future users, remain focused on users' needs, and maintain an empirical foundation for design decisions [36,41,42,54]. Specific implications of designing medication management interventions for the described personas are presented in Table 2.

Table 2. Design implications for Checklist Cheryl, Social Sam, and Responsive Rhonda.

Persona	Unmet needs	Example design implications
Checklist Cheryl, In Control	• No system for tracking changing symptoms and side effects over time • Lacks certainty about when to communicate symptoms and side effects to health care professionals • Medication management is all consuming and leaves little time for self-care	• Adaptive tools to track data (symptoms, side-effects, and strategies) and display trends • AIa-powered database that summarizes academic research • Smartphone apps that combine medication management with self-care for caregivers • Bridge to state and federal caregiver grants
Social Sam, Keeps it Simple	• The current reminder system too simplistic, not accounting for different refill dates • Wants more help and advice from family or friends in his care network	• Text- or email-based reminder systems for reordering different medications • Collaborative platforms for distributing medication management across the care network • Social media platforms, virtual support groups, and online forums for obtaining practical and emotional support from caregiver peers
Responsive Rhonda, Stays Relaxed	• No unmet needs perceived, but may emerge as the care recipient's disease progresses	• Printed educational materials about ADRDb and its progression • Analog tools or infographics that she can situate in her environment • A technology broker who can introduce her to digital medication management supports • Bridge to local aging and disability resource centers

[a]AI: artificial intelligence.

[b]ADRD: Alzheimer disease and related dementias.

The specific attributes of each persona build on the existing literature and extend knowledge in the domain of medication management. In a US-based sample of older adults with chronic heart failure, Holden et al [40] identified the rule-following persona, who has the tendency to respond to uncertainty by seeking reliable rules and clinical expertise. In this study, Checklist Cheryl demonstrates how this tendency extends to medication management processes; Cheryl painstakingly triple checks her medication organization, performs elaborate documentation, and does not trust family or friends to assist in medication management, meaning much of the workload falls on her. Caregivers such as Checklist Cheryl would greatly benefit from state or federal grant programs that pay for respite care or in-home nursing [59]. Such programs require support from the policy makers who set funding priorities and establish budgets. Programs such as these would allow Cheryl to construct a formal support network that she can trust to offset her personal workload.

There are few digital health interventions to support medication management for caregivers, such as Checklist Cheryl, in Control, with some exceptions found internationally [60-62]. As a technology expert, Cheryl may benefit from a multicomponent smartphone app that allows her to, first, view medications and their doses and verify if they have been administered correctly. Second, such an app could enable Checklist Cheryl to document and visualize trends over time, including trends in the care recipient's symptoms, medication side effects, and the effectiveness of strategies [63]. Third, because Cheryl is independent and prefers learning from experts over peers, she may benefit from a feature that uses artificial intelligence to answer medication- and dementia-related questions by synthesizing academic literature [64]. Fourth, Checklist Cheryl needs digital assistance with self-care. A recent international systematic review of mobile apps for family caregivers found only 1 app with a self-care feature [65]. Another systematic review of mobile apps with a caregiver and patient focus found that only 15% of apps included a focus on caregiver self-care [66]. Apps that target self-care needs of other types of busy caregivers (eg, new mothers) could be adapted to fit the self-care needs of caregivers of people living with ADRD [67]. While the typical participant in this study was in their early 60s, younger caregivers are more likely to access the internet and use smartphones and thus may be as interested as Cheryl in digital tools for medication management [68]. Digital tools are also likely to have international applications, as over half of the global population owns a smartphone and about one-third of global smartphone owners have used their phone to access health services [69].

Caregivers such as Social Sam, Keeps it Simple, differ from Checklist Cheryl in that they want simple, rather than complex, medication management systems; deemphasize surveillance in favor of promoting the care recipient's independence; and desire collaboration with other network members as opposed to working alone. While Social Sam's current system is designed for efficiency, it is not sufficiently sophisticated to accommodate polypharmacy, which is increasingly common as the disease progresses [26,70]. Caregivers such as Sam are less technology proficient than those such as Cheryl; he has incorporated the web version of the pharmacy website into his medication acquisition process, but has not used smartphone apps. This may be consistent with other caregivers of Sam's age; research affirms that, though only 61% of those who are aged >65 years own a smartphone, 75% of people in this age range use the internet. Technologically proficient caregivers such as Social

Sam may benefit from websites that support reordering and administering medications, for example, through text- or email-based reminders [71,72].

Social Sam finds his network of family and friends to be helpful, leveraging them to monitor side effects and make clinical judgments. However, similar to other caregivers, those represented by Sam find it challenging to ask others for help [73]. Similar to other male caregivers, in particular, the support network of Sam is small [29]. The current landscape of digital tools for caregivers that support peer interaction is limited [66]. Social Sam would benefit from web-based tools that support distributing caregiving work, and specifically medication management, across a caregiving team [74,75]. Furthermore, Sam would benefit from interventions (eg, virtual support groups and social media platforms) that enable caregivers to connect with peers beyond their geographic area. In the United States, technologies that connect caregivers may have even greater uptake among members of racial and ethnic minority groups, who use social media such as Instagram, Reddit, Snapchat, WhatsApp, and TikTok at higher rates [76]. Reliance on caregiving support from family and friends is also greater among ethnic minority caregivers [77]. More broadly, technologies that connect support networks are needed in low- and middle-income countries, where reliance on unpaid support for dementia caregiving is more common [78].

Caregivers in this study who are represented by Responsive Rhonda, Stays Relaxed, differ from both Checklist Cheryl and Social Sam in that they adopt a fundamentally laid-back management style with little systemization. These caregivers await and respond to medication issues rather than proactively anticipating and addressing them. Similar to the disengaging persona by Holden et al [40], Rhonda does not actively seek out medication information or experiment to improve processes. Responsive Rhonda may represent those caregivers of people in the early stages of dementia, who tend to take fewer medications than those in the later stages [26,27]. As the care recipient's disease progresses, Rhonda would likely benefit from caregiver education and training. Those who are caring for someone in the earlier stages of dementia may be relatively disconnected, with healthcare providers as their sole brokers of support; therefore, health care systems can use personas to create tailored discharge instructions for caregivers such as Responsive Rhonda [79].

While she does not currently report unmet needs related to medication management, as the care recipient's medication regimen becomes more complex, Responsive Rhonda may require a more sophisticated system for medication management. The design of such a system should leverage the tendency of Rhonda to distribute cognition to her physical, rather than digital, environment [80-82]. Rhonda represents caregivers who are novice technology users; therefore, suggestions for Rhonda may also apply to less educated caregivers, who may be less likely and less willing to use digital devices (eg, websites and smartphone apps) to support health, according to a sample of Australian adults [83]. Later in the caregiving journey, caregivers such as Rhonda may benefit from printed materials as well as analog tools (eg, medication organizers) that can be placed as convenient cues throughout the home [80,84]. Rhonda may also benefit from a technology broker who can provide training on using technology to support medication management, for example, by introducing her to specific, informative websites or the mobile app of her preferred pharmacy [85]. This recommendation extends to the 38% of the global population who live in areas in which mobile internet is accessible but who do not currently use it [69].

Limitations and Future Directions

Our study had limitations. While our average participant mirrored the typical US caregiver in many ways [86] and the average care recipient's age matched that of many living with ADRD internationally [87], other sociodemographic characteristics were less representative of caregivers. Our participants had, on average, higher educational attainment, which may result in greater health literacy and enhanced medication management as a result [88]. Furthermore, our sample of caregivers was entirely of White individuals. An analysis of Latino caregivers, for example, would have yielded different personas, as Latino caregivers are more likely to be younger, adult children of care recipients and are less likely to live with the care recipient [86]. One US-based systematic review found that Mexican American, African American, and Korean American caregivers are more committed to keeping care recipients at home rather than long-term care facilities, which may result over time in greater caregiver workload [28]. Moreover, our personas reflect only caregivers who have access to and are willing to participate in research projects, and this group may have unique characteristics (eg, connectedness to academic centers and time to participate in research) [40]. Our personas should not be considered representative of all caregivers managing medication but rather a basis on which future investigations may build [89].

These personas are situated between the conception and gestation stages [36]. That is, while our personas reflect important aspects of our data, have been enriched with storytelling, and deepened through stakeholder involvement, they need further refinement before their full implications can be realized. One limitation of these data is that they reflect only 1 week of caregivers' medication management. A next step in deepening these personas is to collect data about medication management over the course of years, rather than 1 week, to reflect how caregivers' needs change with disease progression. Future work could also incorporate caregivers of people whose cognitive status had been objectively assessed rather than reported by caregivers. While research suggests that family member reports of dementia are often accurate and, when inaccurate, err on the side of underreporting dementia [90,91], it would be important to understand any differences in medication management associated with objectively confirmed dementia diagnoses. Finally, this analysis focused on process- and strategy-related differences in medication management among caregivers. Personas could be further deepened by analyses focused on cognitive, social, or clinical dimensions of medication management [39].

Both our data collection method and the use of personas carry a risk of forgoing the nuance in participant experiences. Our data consist of messages submitted by caregivers and

researchers' follow-up inquiries regarding these messages. When data are collected virtually and participants are asked to provide explanations in retrospect, some degree of detail may be lost (though this is curtailed using messages to aid recall) [92]. Furthermore, in VCI, participants are unlikely to submit messages in inopportune moments, meaning there is a delay not seen in traditional contextual inquiry between performing the task and reporting upon it [46]. While virtual observation may curtail observer effects, it is also possible that participants curate a more flattering self-presentation when communicating digitally versus in person [22]. Last, persona creation inherently emphasizes points of distinction rather than similarities across participants [40]. It is likely that many real caregivers possess qualities of more than one of the personas identified in this study or resemble different personas at different points in their caregiving journey.

Conclusions

This study used a novel methodology, VCI, to understand in rich detail how caregivers manage medications on behalf of those living with ADRD. These data were synthesized into user personas that have direct implications for intervention development. For example, smartphone apps that use artificial intelligence to synthesize information about medication management or meaningfully connect caregivers with others who manage medications could support caregivers who value expertise, such as Checklist Cheryl, or collaboration, such as Social Sam. Before the benefit of medication management personas can be fully realized, these personas must be deepened to reflect nuance and expanded to reflect a more diverse group of caregivers and caregiving experiences. These are vital directions for future research. Once personas have matured in this way, they represent a powerful user-centered strategy to test, evaluate, and disseminate technologies that support medication management among ADRD caregivers.

Acknowledgments

The authors wish to thank Mengwei Tang for her support in formal data analysis. This work was supported by the National Institutes of Health and the National Institute on Aging (grant 1R21AG072418). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Wimo A, Gauthier S. Global estimates of informal care. Alzheimer's Disease International. URL: https://www.alzint.org/u/global-estimates-of-informal-care.pdf [accessed 2024-05-29]
2. Thorpe JM, Thorpe CT, Kennelty KA, Gellad WF, Schulz R. The impact of family caregivers on potentially inappropriate medication use in noninstitutionalized older adults with dementia. Am J Geriatr Pharmacother 2012 Aug;10(4):230-241 [FREE Full text] [doi: 10.1016/j.amjopharm.2012.05.001] [Medline: 22683399]
3. Cotrell V, Wild K, Bader T. Medication management and adherence among cognitively impaired older adults. J Gerontol Soc Work 2006;47(3-4):31-46. [doi: 10.1300/J083v47n03_03] [Medline: 17062521]
4. Fiss T, Thyrian JR, Fendrich K, van den Berg N, Hoffmann W. Cognitive impairment in primary ambulatory health care: pharmacotherapy and the use of potentially inappropriate medicine. Int J Geriatr Psychiatry 2013 Feb;28(2):173-181 [FREE Full text] [doi: 10.1002/gps.3806] [Medline: 22505357]
5. Gench M, Sawan MJ, Langford A, Gnjidic D. Tools to evaluate medication management for caregivers of people living with dementia: a systematic review. Health Expect 2021 Oct;24(5):1570-1581 [FREE Full text] [doi: 10.1111/hex.13318] [Medline: 34289214]
6. Reinhard SC, Levine C, Samis S. Home alone: family caregivers providing complex chronic care. AARP Public Policy Institute. URL: https://www.aarp.org/content/dam/aarp/research/public_policy_institute/health/home-alone-family-caregivers-providing-complex-chronic-care-rev-AARP-ppi-health.pdf [accessed 2024-10-14]
7. While C, Duane F, Beanland C, Koch S. Medication management: the perspectives of people with dementia and family carers. Dementia (London) 2013 Nov;12(6):734-750. [doi: 10.1177/1471301212444056] [Medline: 24337637]
8. Gillespie R, Mullan J, Harrison L. Managing medications: the role of informal caregivers of older adults and people living with dementia. A review of the literature. J Clin Nurs 2014 Dec;23(23-24):3296-3308 [FREE Full text] [doi: 10.1111/jocn.12519] [Medline: 24354583]
9. LaValley SA, Brady LA, Nguyen M, Jacobs D. Family caregivers of older adults: strategies, tools, and technologies for medication management. J Gerontol Soc Work 2023;66(2):291-302 [FREE Full text] [doi: 10.1080/01634372.2022.2100856] [Medline: 35850634]
10. Smith F, Francis SA, Gray N, Denham M, Graffy J. A multi-centre survey among informal carers who manage medication for older care recipients: problems experienced and development of services. Health Soc Care Community 2003 Mar;11(2):138-145 [FREE Full text] [doi: 10.1046/j.1365-2524.2003.00415.x] [Medline: 14629216]

11. Travis SS, Bernard MA, McAuley WJ, Thornton M, Kole T. Development of the family caregiver medication administration hassles scale. Gerontologist 2003 Jun 01;43(3):360-368. [doi: 10.1093/geront/43.3.360] [Medline: 12810899]
12. Smith D, Lovell J, Weller C, Kennedy B, Winbolt M, Young C, et al. A systematic review of medication non-adherence in persons with dementia or cognitive impairment. PLoS One 2017 Feb 6;12(2):e0170651 [FREE Full text] [doi: 10.1371/journal.pone.0170651] [Medline: 28166234]
13. Hawkins LA, Kilian S, Firek A, Kashner TM, Firek CJ, Silvet H. Cognitive impairment and medication adherence in outpatients with heart failure. Heart Lung 2012;41(6):572-582 [FREE Full text] [doi: 10.1016/j.hrtlng.2012.06.001] [Medline: 22784869]
14. Elliott R, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221 [FREE Full text] [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
15. Toot S, Swinson T, Devine M, Challis D, Orrell M. Causes of nursing home placement for older people with dementia: a systematic review and meta-analysis. Int Psychogeriatr 2017 Feb;29(2):195-208 [FREE Full text] [doi: 10.1017/S1041610216001654] [Medline: 27806743]
16. Aston L, Hilton A, Moutela T, Shaw R, Maidment I. Exploring the evidence base for how people with dementia and their informal carers manage their medication in the community: a mixed studies review. BMC Geriatr 2017 Oct 18;17(1):242 [FREE Full text] [doi: 10.1186/s12877-017-0638-6] [Medline: 29047339]
17. Maidment ID, Aston L, Moutela T, Fox CG, Hilton A. A qualitative study exploring medication management in people with dementia living in the community and the potential role of the community pharmacist. Health Expect 2017 Oct;20(5):929-942 [FREE Full text] [doi: 10.1111/hex.12534] [Medline: 28105781]
18. Travis SS, McAuley WJ, Dmochowski J, Bernard MA, Kao HF, Greene R. Factors associated with medication hassles experienced by family caregivers of older adults. Patient Educ Couns 2007 Apr;66(1):51-57 [FREE Full text] [doi: 10.1016/j.pec.2006.10.004] [Medline: 17118616]
19. Look KA, Stone JA. Medication management activities performed by informal caregivers of older adults. Res Social Adm Pharm 2018 May;14(5):418-426 [FREE Full text] [doi: 10.1016/j.sapharm.2017.05.005] [Medline: 28528023]
20. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
21. Evenson S, Muller M, Roth EM. Capturing the context of use to inform system design. J Cogn Eng Decis Mak 2008 Dec 01;2(3):181-203 [FREE Full text] [doi: 10.1518/155534308X377072]
22. Werner NE, Stanislawski B, Marx KA, Watkins DC, Kobayashi M, Kales H, et al. Getting what they need when they need it. Appl Clin Inform 2017 Dec 20;26(01):191-205. [doi: 10.4338/aci-2016-07-ra-0122]
23. Clague F, Mercer SW, McLean G, Reynish E, Guthrie B. Comorbidity and polypharmacy in people with dementia: insights from a large, population-based cross-sectional analysis of primary care data. Age Ageing 2017 Jan 13;46(1):33-39 [FREE Full text] [doi: 10.1093/ageing/afw176] [Medline: 28181629]
24. Mueller C, Molokhia M, Perera G, Veronese N, Stubbs B, Shetty H, et al. Polypharmacy in people with dementia: associations with adverse health outcomes. Exp Gerontol 2018 Jun;106:240-245 [FREE Full text] [doi: 10.1016/j.exger.2018.02.011] [Medline: 29452289]
25. Golchin N, Frank SH, Vince A, Isham L, Meropol SB. Polypharmacy in the elderly. J Res Pharm Pract 2015;4(2):85-88 [FREE Full text] [doi: 10.4103/2279-042X.155755] [Medline: 25984546]
26. Orsel K, Taipale H, Tolppanen AM, Koponen M, Tanskanen A, Tiihonen J, et al. Psychotropic drugs use and psychotropic polypharmacy among persons with Alzheimer's disease. Eur Neuropsychopharmacol 2018 Nov;28(11):1260-1269 [FREE Full text] [doi: 10.1016/j.euroneuro.2018.04.005] [Medline: 30153953]
27. Calvó-Perxas L, Turró-Garriga O, Vilalta-Franch J, Lozano-Gallego M, de Eugenio R, Márquez F, Registry of Dementias of Girona Study Group (ReDeGi Study Group). Trends in the prescription and long-term utilization of antidementia drugs among patients with Alzheimer's disease in Spain: a cohort study using the registry of dementias of Girona. Drugs Aging 2017 Apr;34(4):303-310 [FREE Full text] [doi: 10.1007/s40266-017-0446-x] [Medline: 28258538]
28. Mueller A, Thao L, Condon O, Liebzeit D, Fields B. A Systematic Review of the Needs of Dementia Caregivers Across Care Settings. Home Health Care Manag Pract 2021 Oct 28;34(2):133-147 [FREE Full text] [doi: 10.1177/10848223211056928]
29. Friedman EM, Kennedy DP. Typologies of dementia caregiver support networks: a pilot study. Gerontologist 2021 Nov 15;61(8):1221-1230 [FREE Full text] [doi: 10.1093/geront/gnab013] [Medline: 33585929]
30. Cornet VP, Toscos T, Bolchini D, Rohani Ghahari R, Ahmed R, Daley C, et al. Untold stories in user-centered design of mobile health: practical challenges and strategies learned from the design and evaluation of an app for older adults with heart failure. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17703 [FREE Full text] [doi: 10.2196/17703] [Medline: 32706745]
31. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. 2010. URL: https://www.iso.org/standard/77520.html [accessed 2024-05-29]

32. Mackrill J, Marshall P, Payne SR, Dimitrokali E, Cain R. Using a bespoke situated digital kiosk to encourage user participation in healthcare environment design. Appl Ergon 2017 Mar;59(Pt A):342-356 [FREE Full text] [doi: 10.1016/j.apergo.2016.08.005] [Medline: 27890147]
33. Ganeshkumar P, Bhatnagar A, Burka D, Durgad K, Krishna A, Das B, et al. Discovery, development, and deployment of a user-centered point-of-care digital information system to treat and track hypertension and diabetes patients under India Hypertension Control Initiative 2019-2022, India. Digit Health 2024 May 07;10:20552076241250153 [FREE Full text] [doi: 10.1177/20552076241250153] [Medline: 38715975]
34. Young K, Xiong T, Lee R, Banerjee AT, Leslie M, Ko WY, et al. User-centered design and usability of a culturally adapted virtual survivorship care app for Chinese Canadian prostate cancer survivors: qualitative descriptive study. JMIR Hum Factors 2024 Jan 01;11:e49353 [FREE Full text] [doi: 10.2196/49353] [Medline: 38163295]
35. Tandon A, Cobb B, Centra J, Izmailova E, Manyakov N, McClenahan S, et al. A systematic scoping review of studies describing human factors, human-centered design, and usability of sensor-based digital health technologies. medRxiv Preprint posted online February 24, 2024 [FREE Full text] [doi: 10.1101/2024.02.23.24303220]
36. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design. New York, NY: Elsevier; 2010.
37. McCarthy S, O'Raghallaigh P, Woodworth S, Lim YL, Kenny LC, Adam F. An integrated patient journey mapping tool for embedding quality in healthcare service reform. J Decis Syst 2016 Jun 16;25(sup1):354-368 [FREE Full text] [doi: 10.1080/12460125.2016.1187394]
38. Kiselica AM, Hermann GE, Scullin MK, Benge JF. Technology that CARES: enhancing dementia care through everyday technologies. Alzheimers Dement 2024 Dec;20(12):8969-8978 [FREE Full text] [doi: 10.1002/alz.14192] [Medline: 39508340]
39. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
40. Holden RJ, Daley CN, Mickelson RS, Bolchini D, Toscos T, Cornet VP, et al. Patient decision-making personas: an application of a patient-centered cognitive task analysis (P-CTA). Appl Ergon 2020 Sep;87:103107 [FREE Full text] [doi: 10.1016/j.apergo.2020.103107] [Medline: 32310109]
41. Haldane V, Koh JJ, Srivastava A, Teo KW, Tan YG, Cheng RX, et al. User preferences and persona design for an mHealth intervention to support adherence to cardiovascular disease medication in Singapore: a multi-method study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465 [FREE Full text] [doi: 10.2196/10465] [Medline: 31140445]
42. Engelmann P, Eilerskov N, Thilsing T, Bernardini F, Rasmussen S, Löwe B, et al. Needs of multimorbid heart failure patients and their carers: a qualitative interview study and the creation of personas as a basis for a blended collaborative care intervention. Front Cardiovasc Med 2023;10:1186390 [FREE Full text] [doi: 10.3389/fcvm.2023.1186390] [Medline: 38028443]
43. Werner NE, Campbell NL, Boustani M, Ganci A, Holden RJ. Helping the Helpers - a research protocol for user-centered technology to aid caregiver management of medications for people with Alzheimer's disease and related dementias. Res Social Adm Pharm 2022 Sep;18(9):3680-3686 [FREE Full text] [doi: 10.1016/j.sapharm.2022.03.014] [Medline: 35396151]
44. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340 [FREE Full text] [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
45. Dekker SW, Nyce JM, Hoffman RR. From contextual inquiry to designable futures: what do we need to get there? IEEE Intell Syst 2003 Mar;18(2):74-77 [FREE Full text] [doi: 10.1109/MIS.2003.1193660]
46. Bolger N, Davis A, Rafaeli E. Diary methods: capturing life as it is lived. Annu Rev Psychol 2003 Feb;54(1):579-616. [doi: 10.1146/annurev.psych.54.101601.145030] [Medline: 12499517]
47. Clark RE, Pugh CM, Yates KA, Inaba K, Green DJ, Sullivan ME. The use of cognitive task analysis to improve instructional descriptions of procedures. J Surg Res 2012 Mar;173(1):e37-e42. [doi: 10.1016/j.jss.2011.09.003] [Medline: 22099596]
48. Schlegel EC, Tate JA, Pickler RH, Smith LH. Practical strategies for qualitative inquiry in a virtual world. J Adv Nurs 2021 Oct 04;77(10):4035-4044 [FREE Full text] [doi: 10.1111/jan.15000] [Medline: 34346514]
49. Donnelly T, Rahman-Vyas H, Mandelbaum J, Hatch M, Bhattacharya B. Virtual observation: conducting contextual inquiry remotely. Proc Hum Factors Ergon Soc Annu Meet 2021 Jul 22;10(1):233-237 [FREE Full text] [doi: 10.1177/2327857921101213]
50. Oswald D, Sherratt F, Smith S. Handling the Hawthorne effect: the challenges surrounding a participant observer. Rev Soc Sci 2014;1(1):53-73 [FREE Full text] [doi: 10.21586/ross0000004]
51. Lobo RD, Oliveira MS, Colella JJ, Silva ND, Pastore Junior L, Souza RC. Assessment of the Hawthorne effect during central venous catheter manipulation. Rev Esc Enferm USP 2022;56:e20220125 [FREE Full text] [doi: 10.1590/1980-220X-REEUSP-2022-0125en] [Medline: 36082984]
52. Gaugler JE, Anderson KA, Leach CR, Smith CD, Schmitt FA, Mendiondo M. The emotional ramifications of unmet need in dementia caregiving. Am J Alzheimers Dis Other Demen 2004;19(6):369-380 [FREE Full text] [doi: 10.1177/153331750401900605] [Medline: 15633946]

53. Xhakaj F, Aleven V, McLaren BM. How teachers use data to help students learn: contextual inquiry for the design of a dashboard. In: Proceedings of the 11th European Conference on Technology Enhanced Learning. 2016 Presented at: EC-TEL '16; September 13-16, 2016; Lyon, France p. 340-354 URL: https://link.springer.com/chapter/10.1007/978-3-319-45153-4_26 [doi: 10.1007/978-3-319-45153-4_26]
54. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
55. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023 Oct 25;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
56. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 22;19:160940691989922. [doi: 10.1177/1609406919899220]
57. Varpio L, Ajjawi R, Monrouxe LV, O'Brien BC, Rees CE. Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. Med Educ 2017 Jan 16;51(1):40-50. [doi: 10.1111/medu.13124] [Medline: 27981658]
58. Tracy SJ. Qualitative quality: eight "Big-Tent" criteria for excellent qualitative research. Qual Inq 2010 Oct 01;16(10):837-851. [doi: 10.1177/1077800410383121]
59. How to pay for respite. ARCH National Respite Network and Resource Center. URL: https://archrespite.org/caregiver-resources/how-to-pay-for-respite/ [accessed 2024-10-21]
60. Wagle KC, Skopelja EN, Campbell NL. Caregiver-based interventions to optimize medication safety in vulnerable elderly adults: a systematic evidence-based review. J Am Geriatr Soc 2018 Nov;66(11):2128-2135 [FREE Full text] [doi: 10.1111/jgs.15556] [Medline: 30136714]
61. Kamimura T, Ishiwata R, Inoue T. Medication reminder device for the elderly patients with mild cognitive impairment. Am J Alzheimers Dis Other Demen 2012 Jun;27(4):238-242 [FREE Full text] [doi: 10.1177/1533317512450066] [Medline: 22739031]
62. Shaw CA, Williams KN, Perkhounkova Y, Hein M, Coleman CK. Effects of a video-based intervention on caregiver confidence for managing dementia care challenges: findings from the FamTechCare clinical trial. Clin Gerontol 2020;43(5):508-517 [FREE Full text] [doi: 10.1080/07317115.2020.1729917] [Medline: 32072866]
63. Nie QN, Morrow DG, Azevedo RF, Rogers WA. Developing feedback visualizations to support older adults' medication adherence. Hum Factors Health 2024 Jun;5:100076 [FREE Full text] [doi: 10.1016/j.hfh.2024.100076] [Medline: 40051587]
64. Mostofsky E, Ngo LH, Mittleman MA. The leading AI-powered medical information platform. OpenEvidence. URL: https://openevidence.org [accessed 2024-11-01]
65. Sala-González M, Pérez-Jover V, Guilabert M, Mira JJ. Mobile apps for helping informal caregivers: a systematic review. Int J Environ Res Public Health 2021 Feb 10;18(4):1702 [FREE Full text] [doi: 10.3390/ijerph18041702] [Medline: 33578819]
66. Park JY, Tracy CS, Gray CS. Mobile phone apps for family caregivers: a scoping review and qualitative content analysis. Digit Health 2022 Feb 08;8:20552076221076672 [FREE Full text] [doi: 10.1177/20552076221076672] [Medline: 35154806]
67. Hussain-Shamsy N, Shah A, Vigod SN, Zaheer J, Seto E. Mobile health for perinatal depression and anxiety: scoping review. J Med Internet Res 2020 Apr 13;22(4):e17011 [FREE Full text] [doi: 10.2196/17011] [Medline: 32281939]
68. Faverio M. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-09-27]
69. Shanahan M, Bahia K. The state of mobile internet connectivity 2024. Global System for Mobile Communications Association. URL: https://www.gsma.com/r/somic/ [accessed 2024-09-27]
70. Chen YH, Lai YC, Wu YC, Sasaki J, Tsai KT, Ho CH. Healthcare utilization in different stages among patients with dementia: a nationwide population-based study. Int J Environ Res Public Health 2021 May 26;18(11):5705 [FREE Full text] [doi: 10.3390/ijerph18115705] [Medline: 34073398]
71. Luong P, Glorioso TJ, Grunwald GK, Peterson P, Allen LA, Khanna A, et al. Text message medication adherence reminders automated and delivered at scale across two institutions: testing the nudge system: pilot study. Circ Cardiovasc Qual Outcomes 2021 May;14(5):e007015 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.120.007015] [Medline: 33993727]
72. Gautier JF, Boitard C, Michiels Y, Raymond G, Vergez G, Guedon G. Impact of personalized text messages from pharmacists on medication adherence in type 2 diabetes in France: a real-world, randomized, comparative study. Patient Educ Couns 2021 Sep;104(9):2250-2258. [doi: 10.1016/j.pec.2021.02.022] [Medline: 33750593]
73. Dam AE, Boots LM, van Boxtel MP, Verhey FR, de Vugt ME. A mismatch between supply and demand of social support in dementia care: a qualitative study on the perspectives of spousal caregivers and their social network members. Int Psychogeriatr 2018 Jun;30(6):881-892 [FREE Full text] [doi: 10.1017/S1041610217000898] [Medline: 28606195]
74. Linden A, Jolliff A, Gonzalez D, Loganathar P, Elliott C, Zuraw M, et al. "It made me feel like I wasn't alone in the darkness": exploring dementia care network communication and coordination through a digital health platform. J Am Med Inform Assoc 2022 Nov 14;29(12):2003-2013 [FREE Full text] [doi: 10.1093/jamia/ocac172] [Medline: 36166717]

75. Jolliff A, Coller RJ, Kearney H, Warner G, Feinstein JA, Chui MA, et al. An mHealth design to promote medication safety in children with medical complexity. Appl Clin Inform 2024 Jan 17;15(1):45-54 [FREE Full text] [doi: 10.1055/a-2214-8000] [Medline: 37989249]
76. Social media fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2024-09-27]
77. Pinquart M, Sörensen S. Ethnic differences in stressors, resources, and psychological outcomes of family caregiving: a meta-analysis. Gerontologist 2005 Feb;45(1):90-106. [doi: 10.1093/geront/45.1.90] [Medline: 15695420]
78. Global action plan on the public health response to dementia 2017 - 2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [accessed 2024-04-29]
79. Moore A, Frank C, Chambers LW. Role of the family physician in dementia care. Can Fam Physician 2018 Oct;64(10):717-719 [FREE Full text] [Medline: 30315013]
80. Werner NE, Jolliff AF, Casper G, Martell T, Ponto K. Home is where the head is: a distributed cognition account of personal health information management in the home among those with chronic illness. Ergonomics 2018 Aug;61(8):1065-1078 [FREE Full text] [doi: 10.1080/00140139.2018.1435910] [Medline: 29402181]
81. Mickelson RS, Willis M, Holden RJ. Medication-related cognitive artifacts used by older adults with heart failure. Health Policy Technol 2015 Dec 01;4(4):387-398 [FREE Full text] [doi: 10.1016/j.hlpt.2015.08.009] [Medline: 26855882]
82. Holden RJ, Karanam YL, Cavalcanti LH, Parmar T, Kodthala P, Fowler NR, et al. Health information management practices in informal caregiving: an artifacts analysis and implications for IT design. Int J Med Inform 2018 Dec;120:31-41 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.09.017] [Medline: 30409344]
83. Lee CM, Thomas E, Norman R, Wells L, Shaw T, Nesbitt J, et al. Educational attainment and willingness to use technology for health and to share health information - the reimagining healthcare survey. Int J Med Inform 2022 Aug;164:104803 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104803] [Medline: 35644052]
84. Jolliff AF, Hoonakker P, Ponto K, Tredinnick R, Casper G, Martell T, et al. The desktop, or the top of the desk? The relative usefulness of household features for personal health information management. Appl Ergon 2020 Jan;82:102912 [FREE Full text] [doi: 10.1016/j.apergo.2019.102912] [Medline: 31430599]
85. Hassan AY. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310 [FREE Full text] [doi: 10.2196/20310] [Medline: 32723720]
86. Cothran FA, Alto R, Tomaszewski Farias S, Johnson DK, Lara E, Whitmer RA, et al. Caring for the caregivers: an Alzheimer disease research center call to action. Alzheimer Dis Assoc Disord 2022;36(1):92-95 [FREE Full text] [doi: 10.1097/WAD.0000000000000483] [Medline: 34966023]
87. Dementia fact sheet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2023-09-20]
88. Friis K, Lasgaard M, Rowlands G, Osborne RH, Maindal HT. Health literacy mediates the relationship between educational attainment and health behavior: a Danish population-based study. J Health Commun 2016 Sep 26;21(sup2):54-60. [doi: 10.1080/10810730.2016.1201175] [Medline: 27668691]
89. Young HM, Bell JF, Whitney RL, Ridberg RA, Reed SC, Vitaliano PP. Social determinants of health: underreported heterogeneity in systematic reviews of caregiver interventions. Gerontologist 2020 Feb 14;60(Suppl 1):S14-S28 [FREE Full text] [doi: 10.1093/geront/gnz148] [Medline: 32057083]
90. Cacchione PZ, Powlishta KK, Grant EA, Buckles VD, Morris JC. Accuracy of collateral source reports in very mild to mild dementia of the Alzheimer type. J Am Geriatr Soc 2003 Jun 18;51(6):819-823. [doi: 10.1046/j.1365-2389.2003.51263.x] [Medline: 12757569]
91. Kemp NM, Brodaty H, Pond D, Luscombe G. Diagnosing dementia in primary care: the accuracy of informant reports. Alzheimer Dis Assoc Disord 2002;16(3):171-176. [doi: 10.1097/00002093-200207000-00007] [Medline: 12218648]
92. Kip H, Beerlage-de Jong N, Wentzel J. The contextual inquiry. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. New York, NY: Routledge; 2018:167-186.

Abbreviations

ADRD: Alzheimer disease or related dementias
RQ: research question
UCD: user-centered design
VCI: virtual contextual inquiry

Edited by F Puga; submitted 03.07.24; peer-reviewed by C Laranjeira, FHC Wong, T Wang; comments to author 02.09.24; revised version received 02.11.24; accepted 07.01.25; published 25.07.25.

Please cite as:
Jolliff A, Loganathar P, Holden RJ, Linden A, Patel H, Lee JR, Ganci A, Campbell N, Boustani M, Werner NE
Creating User Personas to Represent the Needs of Dementia Caregivers Who Support Medication Management at Home: Persona Development and Qualitative Study
JMIR Aging 2025;8:e63944
URL: https://aging.jmir.org/2025/1/e63944
doi:10.2196/63944
PMID:

Development and Validation of a Rule-Based Natural Language Processing Algorithm to Identify Falls in Inpatient Records of Older Adults: Retrospective Analysis

Xing Xing Qian[1], PhD; Pui Hing Chau[1], PhD; Daniel Y T Fong[1], PhD; Mandy Ho[1], PhD; Jean Woo[2], MD

[1]School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

[2]Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Pui Hing Chau, PhD
School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

Abstract

Background: In order to address fall underestimation by the *International Classification of Diseases* (*ICD*) in clinical settings, information from clinical notes could be incorporated via natural language processing (NLP) as a possible solution. However, its application to inpatient notes has not been fully investigated.

Objective: This study aims to develop and validate a rule-based NLP algorithm to identify falls based on inpatient admission notes from older patients.

Methods: This retrospective study used 12-year electronic inpatient records of patients aged ≥65 years from public hospitals in Hong Kong. A random sample of 1000 patients was drawn to develop the NLP algorithm. Manual review was the gold standard for assessing the algorithm's performance, with sensitivity, specificity, precision, and F_1-score calculated at the record, episode, and patient levels. In addition, the study compared the number of falls identified by *ICD* codes and clinical notes independently and combined.

Results: Our rule-based NLP algorithm showed excellent performance, with a sensitivity, specificity, precision, and F_1-score of 93.3%, 99.0%, 87.5%, and 0.903 at the record and episode levels, and 92.9%, 98.3%, 89.7%, and 0.912 at the patient level. The combined identification strategy using *ICD* codes and the NLP method provided the most comprehensive capture of fall-related episodes and fallers.

Conclusions: The NLP method proved efficient and accurate in detecting falls from clinical notes in inpatient episodes. For comprehensive capture of fall episodes and fallers, we recommend the combined use of the NLP algorithm and *ICD* codes, which should be applied in future fall epidemiology studies and clinical practice for identifying high-risk groups of fall interventions.

(JMIR Aging 2025;8:e65195) doi:10.2196/65195

KEYWORDS

fall-related admissions; electronic medical records; text mining; case detection; natural language processing

Introduction

Falls are a global public concern and impose considerable health care and economic burdens on the older population [1]. The Decade of Healthy Aging Report from World Health Organization (WHO) highlights fall prevention as an urgent health priority [2]. Fall risk screening is inevitable when identifying vulnerable older individuals for assessment and preventive intervention [3,4]. The history of falls has been widely recognized as a vital component in risk screening [3], given that it doubles the fall risk for older adults [5,6].

In the longitudinal studies on the older population, falls were mainly identified through medical (electronic and nonelectronic) or nonmedical records (eg, retrospective telephone or face-to-face interviews and prospective fall diary or calendar) [6,7]. Compared to nonmedical methods, which are subject to withdrawal or loss of follow-up from participants, medical records are more efficient and cost-saving. Hence, hospitals that provide care for fall-related injuries become the ideal sites to identify and track fallers for interventions. Most studies conducted in clinical settings used the *International Classification of Diseases* (*ICD*) codes to detect falls in the electronic medical records (EMRs) of older adults [6]. Nevertheless, the identification based on the *ICD* codes alone may not completely capture the reported falls. Published studies found that the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes missed 8%

to 20% of falls in patients [8,9]. Involving supplemental records could increase accuracy for fall identification, in which the clinical note review captures the greatest number of events [10]. The traditional manual note review that requires health care professionals as the reviewer is labor-intensive and time-consuming due to the large volume of records. While a relatively expedient and cost-effective approach is searching for the keyword "fall" in the clinical notes, it is important to note that a simple keyword search alone is insufficient to accurately distinguish between fall-related notes and those that do not. This is because notes without actual falls may also contain the term "fall", such as "references to high fall risk," "fall precaution," or "falling asleep," which can result in false positives. To enhance accuracy and overcome this limitation, a novel computer-based approach named natural language processing (NLP) has been applied as an effective solution [11].

NLP algorithms are generally categorized into two approaches: rule based and machine learning (ML) based [11]. While both have their merits, rule-based NLP is selected for this study for several reasons. First, rule-based systems rely on explicitly defined linguistic rules, making them highly interpretable and transparent [12]. This allows developers and users to easily understand and modify the rules, ensuring precise performance. Second, rule-based NLP does not require large annotated datasets, unlike ML-based approaches, which demand extensive training data and computational resource [13]. This makes rule-based systems practical for real-time applications or low-resource environments. Third, rule-based systems efficiently handle edge cases and specific linguistic nuances by directly tailoring rules to unique exceptions—such as distinguishing between actual falls and terms like "falling asleep" in clinical notes—without extensive data or training [12]. While ML-based NLP excels with large-scale datasets and broad contexts [13], it often requires more effort to match the precision and domain-specific adaptability of rule-based systems. In recent years, transformer-based pretrained models have emerged as a powerful ML approach, advancing NLP by capturing long-range dependencies and contextual information in text, even with limited annotated data [14]. Due to their ability to model complex linguistic patterns and contextual relationships, transformer-based models may outperform rule-based systems. Nevertheless, these models still require substantial computational resources, which can be mitigated through cloud-based application programming interfaces (APIs) [15]. In health care systems where patient data confidentiality is strictly restricted, the reliance on internet connectivity to use such models poses a significant challenge. Therefore, for local applications, rule-based approaches remain a pragmatic choice.

However, research on rule-based NLP algorithms for fall identification in clinical notes remains limited. Previous studies have developed rule-based NLP algorithms for inpatient, emergency department (ED), and homecare visit notes [16-18], with the inpatient study specifically focusing on Japanese clinical notes [16]. Although the English-language NLP algorithms for ED and homecare visit notes demonstrated satisfactory performance compared to manual review [17,18], these methods may not be directly transferable to inpatient settings. Hence, more studies are warranted to develop specialized rule-based NLP methods for accurate fall identification in inpatient clinical notes. Furthermore, existing studies primarily compared identification strategies based on single record sources of *ICD* codes or clinical notes [8,9]. There is a lack of studies investigating how the combination of these strategies can enhance fall identification. Therefore, to address these gaps, this study aimed to (1) develop and validate a rule-based NLP algorithm to identify falls based on inpatient admission notes and (2) compare the number of falls identified by independent or combined strategies of *ICD* codes and the NLP method.

Methods

Study Design and Data Source

We performed a retrospective, territory-wide, cohort study using the 12-year (2007-2018) electronic inpatient records of older adults. The dataset used in this study encompasses admissions from all 43 public hospitals of the Hospital Authority (HA), which manage approximately 80% of hospital admissions in Hong Kong and account for nearly 90% of total bed-days [19]. The data were extracted from the Hospital Authority Data Collaboration Lab (HADCL), and the details were described elsewhere [20]. The clinical records were written in English.

Sampling

The sampling frame was a cohort of older adults who (1) resided in Hong Kong, (2) were discharged as an inpatient from public hospitals between January 2007 and December 2017, (3) were aged ≥65 years at admission, (4) had a hospital stay of at least one day for their first inpatient episode, (5) were discharged alive, (6) did not have a hospital discharge from public hospitals 365 days before the first inpatient episode, and (7) were not admitted from the nursing home at their first inpatient episode. More details of the criteria are described elsewhere [20].

To ensure adequate statistical power for developing an NLP algorithm, a sample size of at least 500 is recommended [21]. In this study, a sample of 1000 patients was drawn by random sampling. The electronic clinical notes and diagnosis codes of all inpatient admission episodes from selected patients were extracted for analysis. The dataset contains two types of notes: clinical notes and discharge notes. Only the clinical notes, comprising all textual records associated with linked episodes, were used in this study.

Manual Review

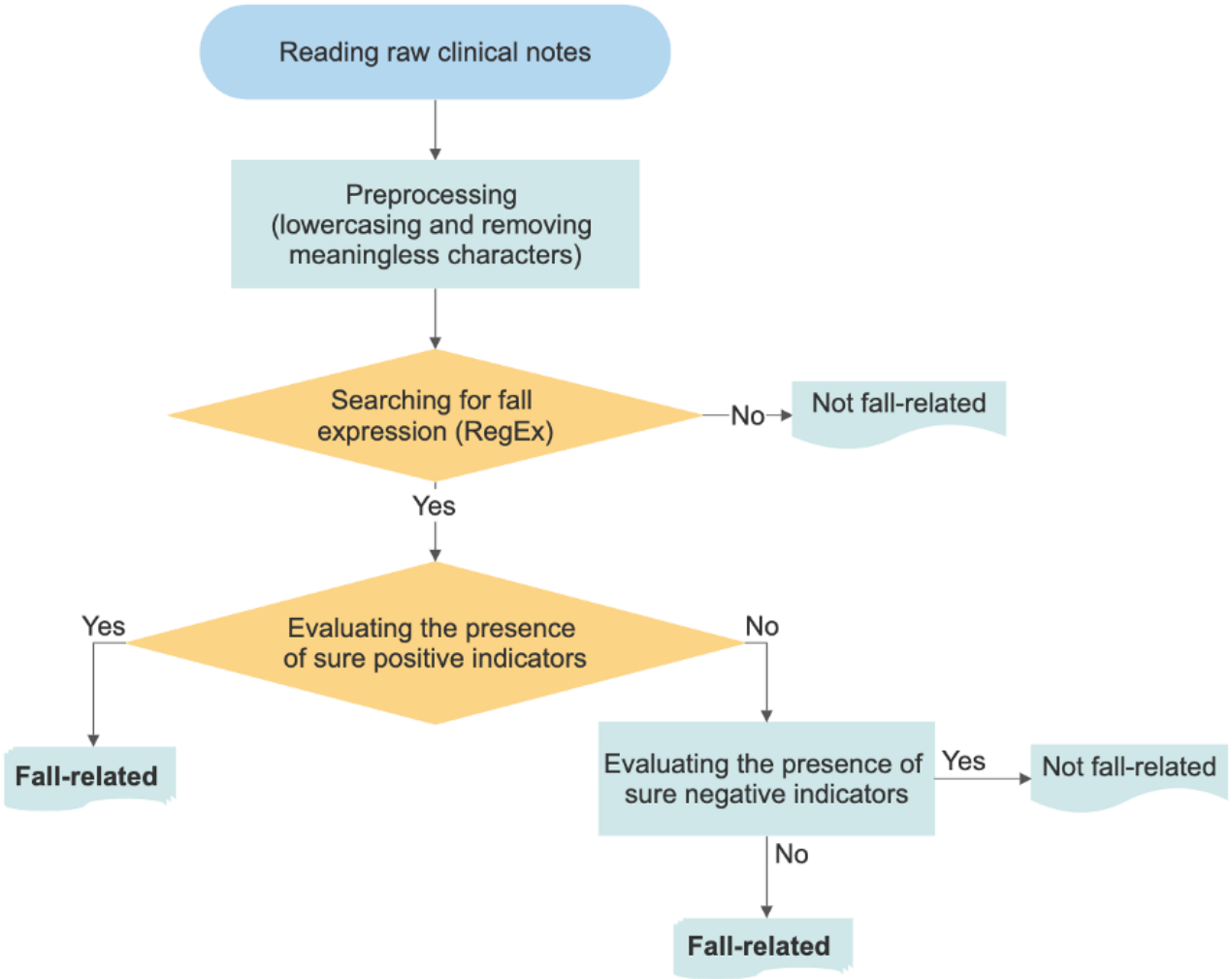
The manual review was considered a gold standard, as consistent with previous studies [16-18]. A total of 2 reviewers independently reviewed each extracted clinical note and determined the corresponding record as being fall related or not, using binary classification. According to the WHO, a fall is defined as an event in which a person unintentionally comes to rest on the ground, floor, or another lower level [1]. An individual who has experienced at least one fall episode is classified as a faller. Reviewers were trained to annotate the clinical notes as fall related if the scenario presented in the notes meet the WHO definition of falls or have specific indicators (eg, "admitted for fall," "presented with fall," "fell onto ground," "slipped and fell," or "fell when..."). Reviewers relied on these

predefined indicators to ensure consistency in their assessments. Any disagreement was resolved by a joint evaluation involving the 2 reviewers and a senior investigator. As each patient could have multiple inpatient admission episodes during the study period and each episode could include multiple records, fall identification was performed at levels of record, episode, and patient. Patients were classified as fallers and episodes were designated as fall related if any associated episodes or records contained falls.

Development and Validation of the Rule-Based NLP Algorithm

Patients were randomly split into the training set for NLP algorithm development and the testing set for validation at a ratio of 80:20 [22]. The development process in the training set included 5 steps (Figure 1). First, clinical notes were preprocessed by lowercasing characters and removing meaningless "Newline Characters" (indicates starting from the next line). Second, the fall expression within the clinical notes was searched by the Regular Expression (RegEx; Section S1 in Multimedia Appendix 1). After filtering out clinical notes with fall expression, the third step evaluated the presence of sure positive indicators (Section S1 in Multimedia Appendix 2) that a fall occurred within the filtered notes. Similarly, the fourth step was evaluating the presence of sure negative indicators (Section S1 in Multimedia Appendix 2) that a fall had not occurred. The sure positive and sure negative indicators were iteratively refined through performance evaluation and rule adjustment. At each cycle, common areas of false positives and negatives were addressed to extend the criteria until the value of sensitivity, specificity, and precision exceeded 90%, and the value of F_1-score exceeded 0.9 at one of the levels for record, episode, or patient [17]. Finally, as fall expression can appear more than once within the same clinical note, the sure positive and negative indicators could coexist in one note. Therefore, clinical notes that (1) had fall expression and (2) had sure positive indicator or did not have sure negative indicator were marked as fall related. Subsequently, the rule-based NLP algorithm was applied to identify falls in clinical notes from the testing set for validation.

Figure 1. Process of the rule-based natural language processing algorithm development. RegEx: Regular Expression.



Comparison of Different Identification Strategies

The number of falls identified by the following methods were compared: (1) manual review of clinical notes, (2) NLP algorithm for clinical notes, (3) *ICD* codes, (4) manual review and *ICD* codes, and (5) NLP algorithm and *ICD* codes. Regarding the strategy of *ICD* codes, a fall was identified by the *ICD-9-CM* E880-E888 for accidental falls or 800 - 909.2,

909.4, 909.9, 910 - 924, 950 - 957, 959 for injury (excluding E000, E800-807, E810-838, E840-858, E860-879, E890-999 for specific external causes of injury) [23,24]. Since not all episodes contain both *ICD* codes and clinical notes—some may lack notes while others may lack *ICD* codes, comprehensive fall identification should consider both data sources. Therefore, the combined strategy, incorporating manual notes review and *ICD* codes, was considered the identification standard to compare the remaining strategies. For combined strategies (4) and (5), the record was identified as fall related when either strategy captured a fall. The comparison was restricted to these 1000 patients, as one of the involved strategies was manual review, which could be time-consuming and labor-intensive if extended to the whole dataset.

Statistical Analysis

The mean and SD values were used to describe the length of clinical notes (number of characters). The falls identified by the developed NLP algorithm were compared with the manual review by confusion matrix to generate standard performance measures of sensitivity, specificity, precision, and F_1-score at the levels of record, episode, and patient in the training and testing sets [25]. Sensitivity refers to the proportion of correctly labeled falls by the NLP algorithm to all falls identified by manual review. Specificity computes the proportion of correctly labeled nonfalls by the NLP algorithm to all nonfalls identified by manual review. Precision reflects the proportion of the correctly labeled falls by the NLP algorithm to all labeled falls by NLP algorithm. F_1-score is the harmonic mean of the precision and sensitivity, with the best value being 1 and the worst being 0. An adequate performance is considered when the value exceeds 90% for sensitivity, specificity, and precision and 0.9 for F_1-score [17]. All data analyses were performed using Python 3.7.6 (CreateSpace).

Ethical Considerations

This study was approved by the Institutional Review Board of the University of Hong Kong and the Hong Kong Hospital Authority West Cluster (approval UW 23 - 046). Informed consent was waived by the Institutional Review Board due to the retrospective, observational nature of the study and the exclusive use of deidentified participant data. Participants were not compensated for their involvement in this study.

Results

In total, 2153 clinical notes (records) of 2095 inpatient admission episodes from 1000 sampled patients were included in the analysis. The mean length of notes was 1072 characters, with an SD of 916. To develop the NLP algorithm, 800 patients (1720 records and 1672 episodes) were randomly assigned to the training set and the remaining 200 patients (433 records and 423 episodes) were assigned to the testing set.

Regarding the training set for development, the manual review identified 140 fall-related records, 135 fall-related episodes, and 127 fallers. In comparison, the NLP algorithm identified 145 fall-related records, 140 fall-related episodes, and 124 fallers. In the testing set for validation, the manual review identified 30 fall-related records, 30 fall-related episodes, and 28 fallers, while the NLP algorithm identified 32 fall-related records, 32 fall-related episodes, and 29 fallers. As shown in Table 1, falls identified by the NLP algorithm were largely concordant with those identified by manual review at record, episode, and patient levels. Performance metrics of the NLP algorithm were satisfactory for both training and testing sets at all 3 levels (Table 2). In the testing set, the NLP algorithm achieved a sensitivity of 93.3%, a specificity of 99.0%, a precision of 87.5%, and an F_1-score of 0.903 at both the record and episode levels. These measures at patient level were 92.9% for sensitivity, 98.3% for specificity, 89.7% for precision, and 0.912 for F_1-score.

Table 3 presents the comparison of different fall identification strategies, with regards to the number of falls identified at the record, episode, and patient levels. Overall, as compared to identification standard of manual notes review combined with *ICD* codes, the NLP algorithm combined with *ICD* codes identified the highest number of falls across all 3 levels (99.4%-104.3%), followed by the NLP algorithm alone (91.1%-93.5%). Specifically, compared to the identification standard, the NLP algorithm combined with *ICD* codes slightly overestimated the numbers of fall-related records and episodes (~4%) yet underestimated the number of fallers (~1%). In contrast, *ICD* codes alone identified the least number of falls (58.7%-60.7%), as 44 inpatient records did not have *ICD-9-CM* codes.

Table . Comparison of fall identification by manual review and the rule-based NLP[a] algorithm.

NLP algorithm and level	Manual review			
	Training set		Testing set	
	Fall, n	Without fall, n	Fall, n	Without fall, n
Record				
Fall	129	16	28	4
Without fall	11	1564	2	399
Episode				
Fall	124	16	28	4
Without fall	11	1521	2	389
Patient				
Fall	116	8	26	3
Without fall	11	665	2	169

[a]NLP: natural language processing.

Table . Performance metrics of the rule-based natural language processing algorithm.

Metric	Training set			Testing set		
	Record level	Episode level	Patient level	Record level	Episode level	Patient level
Sensitivity	92.1%	91.9%	91.3%	93.3%	93.3%	92.9%
Specificity	99%	99%	98.8%	99%	99%	98.3%
Precision	89%	88.6%	93.5%	87.5%	87.5%	89.7%
F_1-score	0.905	0.902	0.924	0.903	0.903	0.912

Table . Number of falls identified by different strategies.

Strategy	Number of falls identified (% of identification standard)		
	Record level	Episode level	Patient level
Manual review and *ICD*[a] codes[b]	193 (100)	184 (100)	168 (100)
NLP[c] algorithm and *ICD* codes	201 (104.1)	192 (104.3)	167 (99.4)
Manual review of clinical notes	170 (88.1)	165 (89.7)	155 (92.3)
NLP algorithm for clinical notes	177 (91.7)	172 (93.5)	153 (91.1)
ICD codes[d]	114 (59.1)	108 (58.7)	102 (60.7)

[a]*ICD*: *International Classification of Diseases*.

[b]Identification standard for comparison.

[c]NLP: natural language processing.

[d]In the whole set, 44 records were without *ICD* codes.

Discussion

Principal Findings

This study developed a rule-based NLP algorithm to identify falls in clinical notes of inpatient admission episodes based on a random, territory-wide sample of older patients, which achieved excellent performance when compared with manual notes review in the testing set. We also compared different fall identification strategies, including the use of *ICD* codes or clinical notes alone or in combination. Our findings provide a feasible and efficient solution for fall screening and shed light on future research and practice related to this field in the clinical setting.

Using NLP algorithms to identify falls can significantly reduce the time and human resources required for the manual notes review. Previous studies have developed NLP methods for fall identification in clinical notes of inpatient wards [16,26,27], and the performance of NLP methods varied. Toyabe [16] derived a rule-based NLP algorithm to detect falls in Japanese progress notes and achieved a sensitivity of 100.0% and a specificity of 98.4%, but with a low precision (6.2%) and an F_1-score of 0.118 at the record level. Shiner et al [26] developed

an ML-based NLP method to capture falls in English progress notes, achieving a relatively low sensitivity of 44.1%, a specificity of 97.0%, a precision of 67.2%, and an F_1-score of 0.533 at the record level. Nakatani et al [27] also developed an ML-based NLP algorithm to differentiate fallers and nonfallers in Japanese nursing notes, with a moderate sensitivity of 76.9% and a specificity of 78.5% at the patient level. In this study's testing set, the sensitivity, specificity, precision, and F_1-score of the rule-based NLP algorithm were 93.3%, 99.0%, 87.5%, and 0.903 at the record level and 92.9%, 98.3%, 89.7%, and 0.912 at the patient level, respectively. The values of these measures were generally higher than previous studies, indicating the promising performance of our rule-based NLP algorithm. Notably, our NLP method was rule based and could be easily applied by researchers or health care professionals with limited programming knowledge to facilitate fall identification in inpatient settings. The rule-based NLP algorithm is also flexible for adjustment based on the different patterns of clinical notes and could be generalized to different systems, such as the ED clinical notes. In addition, future studies may explore the extended application of the NLP method in extracting other fall-related information, such as the circumstance of falls, the mechanism of falls, and resulting injuries, which are important in informing the prognosis and required interventions for patients who have experienced falls [3].

In the epidemiological studies for falls in older adults, the number of fall episodes and fallers are typically the outcomes of interest in estimating the risk (eg, incidence and proportions) and associations. Our findings suggest that the fall identification strategy involving *ICD* codes and clinical notes provides the most comprehensive capture of fall-related episodes and fallers. Identification of falls based on *ICD* codes alone is known to underestimate the number of falls [8,9,28]. Incorporating information from clinical notes is especially beneficial in addressing this issue. Given the limitations of manual review, the NLP algorithm offers an efficient way for applying the combined strategies of clinical notes and *ICD* codes in fall identification on a large scale. Our study found that compared with the combination of manual review and *ICD* codes, the combination of NLP algorithm and *ICD* codes slightly overestimated the number of fall-related episodes (~4%) but not the number of fallers. Therefore, the fall identification strategy combining NLP algorithm and *ICD* codes with feasibility and accuracy has promising application value in risk screening and patient tracking.

The health care workforce shortage is a global problem. Identification of efficient methods to enhance the quality of patient care and safety is a priority agenda in health care field. In terms of implications to practice, there are posthospital health care services or fall prevention programs designed to improve functions and reduce falls for patients recently discharged from hospitals. Fall history is a dominant risk factor for posthospital falls in older patients [6]. Since hospitals are ideal sites for identifying and tracking patients with a history of falls, this combined strategy of NLP method and *ICD* codes is expected to benefit posthospital fall interventions both at baseline and during follow-up. The combined strategy allows for standardized and consistent screening across an extensive EMR dataset, and it can be integrated into existing EMR systems to conduct real-time screening and identification of fall events within minutes. For instance, upon discharge, health care professionals can use it to screen older patients with a history of falls to be recruited in posthospital fall prevention programs. After discharge, patients' fall-related readmission can be tracked and identified timely by this strategy in the EMR system. To establish a more comprehensive surveillance procedure, supplementing nonmedical methods to collect fall incidents that do not present hospitals can be considered. Meanwhile, posthospital health care services, such as the Nurse and Allied Health Clinic (NAHC) for Falls, are designed specifically to prevent falls for patients. Especially for a resource-limited setting, such services can better target patients with a fall history.

The strengths of this study include using 12-year medical records from a territory-wide older population across multiple hospitals, which enhances the reliability and generalizability of the developed NLP algorithm. In addition, we compared the performance of independent and combined fall identification strategies based on *ICD* codes and clinical notes to determine the optimal method in practice. However, several limitations should be acknowledged. First, direct comparison of our NLP algorithm with existing rule-based and ML-based NLP algorithms, including transformer-based pretrained models, was not conducted in this study. Future research should compare the performance of these approaches on the same set of inpatient notes and evaluate their strengths and limitations. Second, a comprehensive analysis of false positives and negatives pattern could not be presented due to data sharing restrictions. Third, due to data availability, this study was limited by the inability to incorporate clinical notes from patients admitted to the ED to derive and compare NLP algorithms for different health care utilizations. Future studies are warranted to compare NLP methods generated from different settings and evaluate their contribution to fall identification. In addition, since the data in this study were drawn from 12-year medical records, it is uncertain whether the writing style of doctors or nurses changed over time. The lack of interrater reliability testing represents another limitation of this study. While the analysis of performance metrics for *ICD* codes would provide a more comprehensive comparison, this analysis was not conducted in this study and remains an important direction for future research. Finally, this study used a cohort developed for another study, which excluded older patients who had hospitalizations within the past year or were admitted from nursing homes, leading to a relatively healthier sample with lower fall risk. Investigations on the performance of our rule-based NLP algorithm in a more general older population are required.

Conclusions

To conclude, this study has demonstrated that the rule-based NLP algorithm could efficiently and accurately detect falls in clinical notes from inpatient admission episodes with excellent performance. The combined identification strategy that incorporates the NLP algorithm and *ICD* codes could be applied to comprehensively capture fall episodes and fallers in future research of fall epidemiology, as well as in clinical practice to facilitate the identification of high-risk groups for fall interventions.

Acknowledgments

The authors would like to thank the Hong Kong Hospital Authority and its Information Technology and Health Informatics Department for providing the data and supporting this study.

Data Availability

The datasets generated during and analyzed during this study are not publicly available as its owned by the Hong Kong Hospital Authority.

Authors' Contributions

XXQ, PHC, DYTF, MH, and JW contributed to the study concept and design; XXQ and PHC contributed to the acquisition, analysis, and interpretation of data; XXQ contributed to the drafting of the manuscript; XXQ, PHC, DYTF, MH, and JW contributed to critical revision of the manuscript for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Regular expressions used to identify fall events.
[DOCX File, 34 KB - aging_v8i1e65195_app1.docx]

Multimedia Appendix 2
Criteria for sure positive and sure negative fall indicators.
[DOCX File, 37 KB - aging_v8i1e65195_app2.docx]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2024-04-19]
2. Decade of healthy ageing: baseline report. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240017900 [accessed 2024-04-19]
3. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
4. Cameron ID, Dyer SM, Panagoda CE, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018 Sep 7;9(9):CD005465. [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
5. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. Epidemiology 2010 Sep;21(5):658-668. [doi: 10.1097/EDE.0b013e3181e89905] [Medline: 20585256]
6. Qian XX, Chen Z, Fong DYT, Ho M, Chau PH. Post-hospital falls incidence and risk factors among older adults: a systematic review and meta-analysis. Age Ageing 2022 Jan 6;51(1):afab209. [doi: 10.1093/ageing/afab209]
7. Schwenk M, Lauenroth A, Stock C, et al. Definitions and methods of measuring and reporting on injurious falls in randomised controlled fall prevention trials: a systematic review. BMC Med Res Methodol 2012 Apr 17;12:50. [doi: 10.1186/1471-2288-12-50] [Medline: 22510239]
8. Patterson BW, Smith MA, Repplinger MD, et al. Using chief complaint in addition to diagnosis codes to identify falls in the emergency department. J Am Geriatr Soc 2017 Sep;65(9):E135-E140. [doi: 10.1111/jgs.14982] [Medline: 28636072]
9. Patterson BW, Jacobsohn GC, Maru AP, et al. RESEARCHComparing strategies for identifying falls in older adult emergency department visits using EHR data. J Am Geriatr Soc 2020 Dec;68(12):2965-2967. [doi: 10.1111/jgs.16831] [Medline: 32951200]
10. Olsen S, Neale G, Schwab K, et al. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. Qual Saf Health Care 2007 Feb;16(1):40-44. [doi: 10.1136/qshc.2005.017616] [Medline: 17301203]
11. Trinh VQN, Zhang S, Kovoor J, et al. The use of natural language processing in detecting and predicting falls within the healthcare setting: a systematic review. Int J Qual Health Care 2023 Oct 17;35(4):mzad077. [doi: 10.1093/intqhc/mzad077] [Medline: 37758209]
12. Liu H, Cocea M. Fuzzy rule based systems for interpretable sentiment analysis. Presented at: 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI); Feb 4-6, 2017; Doha, Qatar p. 129-136. [doi: 10.1109/ICACI.2017.7974497]

13. Spasic I, Nenadic G. Clinical text data in machine learning: systematic review. JMIR Med Inform 2020 Mar 31;8(3):e17984. [doi: 10.2196/17984] [Medline: 32229465]
14. Rahali A, Akhloufi MA. End-to-end transformer-based models in textual-based NLP. AI 2023;4(1):54-110. [doi: 10.3390/ai4010004]
15. Yenduri G, Ramalingam M, Selvi GC, et al. GPT (Generative Pre-Trained Transformer)— a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. IEEE Access 2024;12:54608-54649. [doi: 10.1109/ACCESS.2024.3389497]
16. Toyabe SI. Detecting inpatient falls by using natural language processing of electronic medical records. BMC Health Serv Res 2012 Dec 5;12:448. [doi: 10.1186/1472-6963-12-448] [Medline: 23217016]
17. Patterson BW, Jacobsohn GC, Shah MN, et al. Development and validation of a pragmatic natural language processing approach to identifying falls in older adults in the emergency department. BMC Med Inform Decis Mak 2019 Jul 22;19(1):138. [doi: 10.1186/s12911-019-0843-7] [Medline: 31331322]
18. Topaz M, Murga L, Gaddis KM, et al. Mining fall-related information in clinical notes: Comparison of rule-based and novel word embedding-based machine learning approaches. J Biomed Inform 2019 Feb;90:103103. [doi: 10.1016/j.jbi.2019.103103] [Medline: 30639392]
19. Report of the strategic review on healthcare manpower planning and professional development. Health Bureau of the Government of the Hong Kong Special Administrative Region. 2020. URL: https://www.healthbureau.gov.hk/en/press_and_publications/otherinfo/180500_sr/srreport.html [accessed 2024-04-19]
20. Qian XX, Chau PH, Fong DYT, Ho M, Woo J. Post-hospital falls among the older population: the temporal pattern in risk and healthcare burden. J Am Med Dir Assoc 2023 Oct;24(10):1478-1483. [doi: 10.1016/j.jamda.2023.07.014] [Medline: 37591487]
21. Juckett D. A method for determining the number of documents needed for a gold standard corpus. J Biomed Inform 2012 Jun;45(3):460-470. [doi: 10.1016/j.jbi.2011.12.010] [Medline: 22245601]
22. Gholamy A, Kreinovich V, Kosheleva O. Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation. Int J Intell Technol Appl Stat 2018;11(2):105-111. [doi: 10.6148/IJITAS.201806_11(2).0003]
23. Lam C, Kang JH, Lin HY, Huang HC, Wu CC, Chen PL. First fall-related injuries requiring hospitalization increase the risk of recurrent injurious falls: a nationwide cohort study in Taiwan. PLoS One 2016;11(2):e0149887. [doi: 10.1371/journal.pone.0149887] [Medline: 26900917]
24. Owens PL, Russo A, Spector W, Mutter R. Emergency department visits for injurious falls among the elderly, 2006. Agency for Healthcare Research and Quality. 2009. URL: https://hcup-us.ahrq.gov/reports/statbriefs/sb80.pdf [accessed 2025-06-06]
25. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manag 2009 Jul;45(4):427-437. [doi: 10.1016/j.ipm.2009.03.002]
26. Shiner B, Neily J, Mills PD, Watts BV. Identification of inpatient falls using automated review of text-based medical records. J Patient Saf 2020 Sep;16(3):e174-e178. [doi: 10.1097/PTS.0000000000000275] [Medline: 27331601]
27. Nakatani H, Nakao M, Uchiyama H, Toyoshiba H, Ochiai C. Predicting inpatient falls using natural language processing of nursing records obtained from Japanese electronic medical records: case-control study. JMIR Med Inform 2020 Apr 22;8(4):e16970. [doi: 10.2196/16970] [Medline: 32319959]
28. Waters TM, Chandler AM, Mion LC, et al. Use of International Classification of Diseases, Ninth Revision, Clinical Modification, codes to identify inpatient fall-related injuries. J Am Geriatr Soc 2013 Dec;61(12):2186-2191. [doi: 10.1111/jgs.12539] [Medline: 24329820]

Abbreviations

API: application programming interface
ED: emergency department
EMR: electronic medical record
HA: Hospital Authority
HADCL: Hospital Authority Data Collaboration Lab
***ICD*:** *International Classification of Diseases*
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
ML: machine learning
NAHC: Nurse and Allied Health Clinic
NLP: natural language processing
RegEx: Regular Expression
WHO: World Health Organization

Edited by M O'Connell; submitted 07.08.24; peer-reviewed by A Mitra, H Wang, K Zandbiglari, S Zeng; revised version received 25.03.25; accepted 10.05.25; published 08.07.25.

Please cite as:
Qian XX, Chau PH, Fong DYT, Ho M, Woo J
Development and Validation of a Rule-Based Natural Language Processing Algorithm to Identify Falls in Inpatient Records of Older Adults: Retrospective Analysis
JMIR Aging 2025;8:e65195
URL: https://aging.jmir.org/2025/1/e65195
doi:10.2196/65195

The Impact of Vision Impairment on Self-Reported Falls Among Older US Adults: Cross-Sectional and Longitudinal Study

Kasem Seresirikachorn[1,2], MD, PhD; Rachasak Somyanonthanakul[3], PhD; Matthew Johnson[1], BS; Panisa Singhanetr[1,4], MD; Jiraporn Gatedee[5], PhD; David Friedman[1], MD, MPH, PhD; Nazlee Zebardast[1], MSc, MD

[1]Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, MA, United States

[2]Department of Ophthalmology, College of Medicine, Rangsit University, Rajavithi Hospital, Bangkok, Thailand

[3]Data Science and Innovation, College of Interdisciplinary Studies, Thammasat University, Pathum Thani, Thailand

[4]Mettapracharak Eye Institute, Mettapracharak (Wat Rai Khing) Hospital, Nakhon Pathom, Thailand

[5]Faculty of Medical Technology, Rangsit University, Pathum Thani, Thailand

Corresponding Author:
Nazlee Zebardast, MSc, MD
Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, MA, United States

Abstract

Background: Falls are the leading cause of injury among older adults, with vision impairment recognized as a significant risk factor. However, many existing studies have been limited by small sample sizes, retrospective designs, or insufficient adjustment for confounding factors. To overcome these limitations, we used data from the University of Michigan's Health and Retirement Study (HRS) to analyze the association between self-reported vision and fall risk among older adults in a large, nationally representative sample.

Objective: The objective of this study was to investigate the association between vision impairment and falls and assess whether subjective vision impairment predicts future falls in older adults.

Methods: This cross-sectional and longitudinal analysis used data from the HRS (1996 - 2020) to assess the relationship between self-reported vision, glaucoma history, and falls among US adults aged 65 years and older. HRS uses a biennial, multistage area probability sample survey design, collecting data with community-dwelling individuals followed up every 2 years until death, tracking health, economic, and social outcomes. Multivariate logistic regression was used to analyze associations between self-reported vision and self-reported falls in the past 2 years.

Results: A total of 38,835 respondents contributed 117,834 observations. The weighted proportion of participants reporting falls was 37.9% (95% CI 37.7% - 40.1%). Significant risk factors for falls included overall eyesight impairment (adjusted odds ratio [aOR] 1.36, 95% CI 1.20 - 1.56), distance vision impairment (aOR 1.37, 95% CI 1.32 - 1.42), near vision impairment (aOR 1.33, 95% CI 1.27 - 1.37), and glaucoma (aOR 1.15, 95% CI 1.07 - 1.24). A similar association was observed for serious falls, where overall eyesight impairment (aOR 1.20, 95% CI 1.03 - 1.44), distance vision impairment (aOR 1.14, 95% CI 1.07 - 1.22), near vision impairment (aOR 1.12, 95% CI 1.05 - 1.18), and glaucoma (aOR 1.15, 95% CI 1.05 - 1.26) were significant. In longitudinal analyses, overall vision impairment (aOR 1.23, 95% CI 1.16 - 1.29), distance vision impairment (aOR 1.27, 95% CI 1.20 - 1.38), near vision impairment (aOR 1.23, 95% CI 1.19 - 1.32), and glaucoma (aOR 1.25, 95% CI 1.13 - 1.37) increased the risk of future falls. Reported overall vision was significantly associated with the number of falls in both the same (P<.001) and subsequent (P<.001) survey cycles.

Conclusions: Both distance and near vision impairment, as well as glaucoma, are associated with a higher risk of falls in older adults and present possible areas for intervention and prevention.

(JMIR Aging 2025;8:e68771) doi:10.2196/68771

KEYWORDS

elderly; falls; glaucoma; Health and Retirement Study; vision impairment

Introduction

Falls are the leading cause of injury among adults ages 65 years or older [1]. Nearly 30% of older adults fall each year, and 10% report injury from a fall [2]. In the United States in 2018, there were 36 million reported falls, resulting in 3 million emergency department visits, nearly 1 million hospitalizations, and 32,000 deaths [3]. Approximately 1 in 10 falls in older adults results in serious injury, such as hip fracture or traumatic brain injury, and the annual cost of falls in the United States is more than US $50 billion [4-6]. Globally, falls are the 13th leading cause of death, and the prevalence of falls is increasing as the

population gets older [7]. Fortunately, multifactorial intervention strategies such as medical staff and patient education, providing supplies for hip protection, and sensory optimization by providing hearing aids have all been able to successfully reduce the risk of falling [5,7,8]. Home modifications to improve safety, such as changing floor coverings, removing loose carpets, fitting handrails, maintaining steps and ramps, and improving lighting, can reduce falling hazards in homes.

Falls are significant, life-altering health events with modifiable risk factors such as impaired balance, polypharmacy, and poor vision [9-11]. Impaired vision has been shown to be a risk factor for falls in older adults. In 1998, the Beaver Dam Eye Study showed that impaired visual acuity was associated with an increased risk of falling, and older adults in the worst visual acuity group were significantly more likely to have fallen 2 or more times in the past year [12]. The Blue Mountains Eye Study similarly found a significant relationship between low visual acuity and falls. In addition, they reported that decreased contrast sensitivity and reduced visual field (VF) increased the risk of falls and proposed that the presence of a cataract may explain this association [13]. In the Salisbury Eye Evaluation study, only reduced peripheral VF was associated with the risk of falling, while visual acuity, contrast sensitivity, and stereo acuity were not after adjusting for potential confounders [14]. Other studies support a relationship between impaired vision and falls but differ as to which components of visual function increase fall risk [5,9,15-17].

These studies provided valuable information that broadly supports impaired vision as a modifiable risk factor for falls; however, many of these studies had limitations due to small sample sizes, retrospective analyses, or the inability to adjust for confounders. Even the most extensive studies to date have data spanning only 1 - 2 years and are generally limited to local communities or geographic regions [12-14].

To address these limitations, we used the University of Michigan Health and Retirement Study (HRS) data. The HRS is a diverse, nationally representative longitudinal panel study of US adults ages 50 years and older, followed every 2 years for over 20 years [18]. Data on fall incidence and self-reported vision impairment were collected from over 25,000 respondents over a period of two decades. [7]. Self-reported vision has been established as a practical proxy measure for visual function, demonstrating a strong relationship with various objective visual function assessments [19,20]. This indicates that self-reported vision can serve as a reliable indicator of visual acuity. Furthermore, self-reported vision assessments offer a valuable advantage in under-resourced areas, where accessible and convenient tools for vision evaluation are scarce. By leveraging self-reported vision data, researchers and health care professionals can overcome logistical barriers and better understand vision-related needs in these communities. A better understanding of what aspects of vision impairment increase fall risk will help clarify promising vision interventions to decrease the rate of falls among older adults.

The objective of this study was to investigate the relationship between various aspects of self-reported vision impairment and falls and prospectively assess whether subjective vision impairment is predictive of future falls in a large and diverse longitudinal population-based study of older adults.

Methods

Data Collection

Data were from the HRS database, a longitudinal panel study that surveys a nationally representative sample of Americans aged 50 years and older every two years, with follow-up continuing until death. The HRS uses a biennial, multistage area probability sample survey design, using a combination of in-person and telephone interviews to collect data. The survey focuses on community-dwelling individuals. This longitudinal design enables the tracking of participants' health, economic, and social outcomes over time, providing a comprehensive understanding of aging and retirement in America. The HRS is supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration. For this study, we used data from 14 waves of the HRS (1996 - 2020).

The analytic subsample for this study comprises community-dwelling respondents aged 65 years and older from the 1996 - 2020 HRS survey rounds, as the falls questionnaire was only administered to participants at or older than this age threshold, with all included variables having item nonresponse rates below 10%.

Ethical Considerations

Since the HRS database is deidentified and publicly accessible, this study was exempt from the Mass General Brigham institutional review board approval and patient permission was waived. The study adhered to the tenets of the Declaration of Helsinki.

Primary Outcomes and Definitions of Variables

The primary outcome measures included self-reported falls, serious injury falls, and the number of falls. Participants were asked if they had fallen in the past two years. If they answered "yes," they were then asked how many times they had fallen and whether any of the falls were serious. Serious injury falls were defined as those requiring medical treatment.

Ophthalmic parameters were acquired from the HRS survey assessments and included overall eyesight, distance vision, and near vision (see Multimedia Appendix 1), each rated on a scale from 1 to 5, with scores of 4 (fair) and 5 (poor) considered indicative of impaired vision. For overall eyesight, a score of 6 could be assigned for patients with legal blindness. The history of glaucoma was based on whether participants reported ever being treated for glaucoma.

The HRS also collected information on age, gender, race, level of education, marital status, self-reported diagnoses of hypertension, diabetes, stroke, heart conditions (including heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems), arthritis or rheumatism, smoking and alcohol consumption status, psychiatric medication use, exercise activities, and household income.

Study Timepoints of Data Analysis

The data were analyzed cross-sectionally and longitudinally to confirm the correlations between ophthalmic parameters and falls. Each participant observation was analyzed as individual encounters in both analyses. In the cross-sectional analysis, self-reported ophthalmic parameters and falls were ascertained from the same 2-year interview cycle. Each completed survey was counted as 1 entry. In the longitudinal analysis, the self-reported ophthalmic parameters were obtained during the survey and analyzed with the reported falls recorded in the patient's survey 2 years later.

Statistical Analysis

Data are presented as base-year weighted proportions, accounting for the Health and Retirement Study adjustments for attrition and sampling error. Descriptive statistics were used to summarize participant data on falls outcomes, including falls, no falls, serious falls, and nonserious falls. We used *t* tests for continuous variables and χ^2 tests for categorical variables to evaluate differences between outcomes. We used univariable and multivariable logistic regression models to determine the cross-sectional and longitudinal association between ophthalmic parameters and falls, adjusting for age and other potential confounders by calculating the odds ratio (OR) and their corresponding 95% CI. For the cross-sectional model, we assessed the association between vision status and reported falls in the same 2-year survey cycle, while for the longitudinal model, we considered falls reported in the subsequent 2-year survey cycle.

All variables with a *P* value of <.05 in the univariable analyses (including age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol status, psychiatric medications, degree of exercise, and household income) were included in multivariable logistic regression models. Separate multivariable regression models were used to determine the relationship between each ophthalmic parameter and falls and serious falls. In addition, the association between the degree of vision impairment and the number of falls was assessed using a 1-way analysis of variance. All models were analyzed using generalized estimating equations for multiple observations from the same participant and multiple imputation techniques to replace the missing data. Data analysis was done with IBM SPSS Statistics 22.0 for Microsoft Windows and *P* values <.05 were considered statistically significant.

Results

Demographics of Study Sample

A total of 38,835 respondents contributed to 117,834 participant-observations (see Tables 1 and 2). The total number of observations in which participants reported a total of 40,477 falls and 13,471 serious falls. The weighted proportion of participants who reported any falls was OR 37.9% (95% CI 37.7% - 40.1%), while the proportion reporting serious falls was OR 30.9% (95% CI 30.7% - 31.1%). Furthermore, the weighted proportion of participants who experienced falls in the future was OR 36.5% (95% CI 35.8% - 37.2%), and OR 32.5% (95% CI 32.2% - 32.9%) for future serious falls. Falls increased with older age, and over half of those aged 90 years and older reported falling. The proportion of falls qualifying as serious also increased with age. White people had the highest fall rate, and a higher proportion of falls and serious falls was observed in females and individuals with lower levels of education. In total, 50% of patients with a history of stroke reported falling, with one-third of the falls being serious falls.

Table . Baseline characteristics of study participants comparing those who reported falls versus no falls (total number of participant-observations).

Characteristic	Participants with any self-reported falls (n=40,477)	Participants with no self-reported falls (n=77,357)	*P* value
Total (%), OR[a] (95% CI)	37.9 (37.7 - 40.1)	62.1 (61.9 - 62.3)	—[b]
Age in years (%), OR (95% CI)			<.001
65 - 69	28.5 (28.2 - 28.7)	71.5 (71.2 - 71.9)	
70 - 74	30.9 (30.5 - 31.4)	69.1 (68.6 - 69.7)	
75 - 79	32.9 (32.3 - 33.7)	67.1 (66.5-67.8)	
80 - 84	37.9 (37.5 - 38.5)	62.1 (61.6-62.9)	
85 - 89	44.8 (44.2 - 45.6)	55.2 (54.8-55.9)	
90+	52.2 (51.6-52.7)	47.8 (46.8-48.4)	
Sex (%), OR (95% CI)			<.001
Male	30.2 (29.8 - 30.8)	69.8 (69.2 - 70.3)	
Female	35.7 (35.3 - 36.1)	64.3 (63.8 - 64.8)	
Ethnicity (%), OR (95% CI)			<.001
Hispanic	33.1 (31.8 - 33.9)	66.9 (65.8 - 67.9)	
Non-Hispanic	34.7 (34.1 - 35.3)	65.3 (65.0 - 65.8)	
Race (%), OR (95% CI)			<.001
White	34.7 (34.1 - 35.3)	65.3 (64.3 - 66.4)	
Black or African American	28.9 (28.0 - 29.9)	71.1 (70.1 - 72.0)	
Others[c]	30.7 (29.9 - 31.8)	69.3 (68.6 - 70.1)	
Education (%), OR (95% CI)			<.001
<High school	35.1 (34.6 - 35.7)	64.9 (63.8 - 65.5)	
High school degree	33.2 (32.8 - 33.5)	66.8 (65.8 - 67.2)	
>High school degree	31.7 (31.1 - 32.3)	68.3 (67.8 - 68.8)	
Marital status (%), OR (95% CI)			<.001
Married	61.0 (60.8-61.2)	39.0 (38.9-39.1)	
Widowed, separated, or divorced	59.3 (59.0-59.8)	40.7 (40.2-41.3)	
Never married	38.7 (38.2-39.2)	61.3 (60.9-61.7)	
Comorbidities (%), OR (95% CI)			
Hypertension	35.7 (35.2-36.1)	64.3 (63.5-64.8)	<.001
Diabetes mellitus	39.7 (39.4-40.2)	60.3 (59.8-60.7)	<.001
Heart	42.0 (41.6-42.5)	58.0 (57.5-58.5)	<.001
Stroke	49.2 (48.3-50.0)	50.8 (50.3-51.5)	<.001
Arthritis (%), OR (95% CI)	36.8 (35.9 - 38.0)	63.2 (61.9 - 64.1)	<.001
Current smoking (%), OR (95% CI)	45.0 (44.3 - 46.0)	56.0 (54.9 - 57.1)	<.001
Active alcohol drinking (%), OR (95% CI)	55.2 (54.8 - 55.7)	44.8 (44.0 - 45.7)	<.001
Psychiatric medication (%), OR (95% CI)	62.1 (61.5 - 62.8)	37.9 (37.1 - 38.6)	<.001
Vigorous exercise (%), OR (95% CI)	70.6 (70.0 - 71.2)	29.4 (28.8 - 29.7)	<.001
Household income (US $), mean (95% CI)	64,319.1 (56,172.2 - 71,816.3)	69,983.7 (45,562.1 - 73,805.4)	<.001

[a]OR: odds ratio.

[b]Not applicable.

[c]Other races: Asian, Latino, and Native American.

Table . Baseline characteristics of study participants comparing those who reported serious falls versus no serious falls (total number of participant-observations).

Characteristics	Participants with any self-reported serious falls (n=13,471)	Participants with no self-reported serious falls (n=27,006)	*P* value
Total (%), OR[a] (95% CI)	30.9 (30.7 - 31.1)	69.1 (68.9 - 69.3)	—[b]
Age in years (%), OR (95% CI)			<.001
65-69	21.4 (21.3 - 21.5)	78.6 (78.3-78.8)	
70-74	31.2 (31.0-31.4)	68.8 (68.4-69.1)	
75-79	40.6 (40.2-40.9)	59.4 (59.0-59.8)	
80-84	42.6 (42.1-43.0)	57.4 (56.9-57.9)	
85-89	43.9 (43.0-44.7)	56.1 (55.7-56.6)	
90+	39.9 (39.3-40.6)	60.1 (59.1-61.1)	
Sex (%), OR (95% CI)			<.001
Male	23.8 (23.7 - 23.9)	76.2 (75.9 - 76.5)	
Female	39.1 (38.9-39.4)	60.9 (60.3-61.5)	
Ethnicity (%), OR (95% CI)			<.001
Hispanic	26.9 (26.3 - 27.5)	73.1 (72.4 - 73.9)	
Non-Hispanic	31.8 (31.2 - 32.4)	68.2 (67.6 - 68.8)	
Race (%), OR (95% CI)			<.001
White	35.8 (35.2 - 36.3)	64.2 (63.7 - 64.8)	
Black or African American	20.0 (19.6 - 20.5)	80.0 (79.7 - 80.4)	
Others[c]	19.9 (19.1 - 20.8)	80.1 (79.6 - 80.9)	
Education (%), OR (95% CI)			<.001
<High school	31.9 (31.3 - 32.7)	68.1 (67.5 - 68.6)	
High school degree	34.1 (33.7 - 34.5)	65.9 (65.5 - 66.6)	
>High school degree	26.9 (26.2 - 27.4)	3.1 (72.7 - 73.6)	
Marital status (%), OR (95% CI)			<.001
Married	31.8 (31.3 - 32.3)	68.2 (67.9 - 68.6)	
Widowed, separated, or divorced	32.8 (32.3 - 33.3)	67.2 (66.8 - 67.9)	
Never married	17.8 (17.3 - 18.2)	82.2 (81.8 - 82.6)	
Comorbidities (%), OR (95% CI)			
Hypertension	32.8 (32.2-33.7)	67.2 (66.9-67.6)	<.001
Diabetes mellitus	32.3 (31.9-33.6)	67.7 (67.0-68.3)	<.001
Heart	35.0 (34.4-35.8)	65.0 (64.6-65.8)	<.001
Stroke	35.6 (34.5-36.3)	64.4 (64.1-64.7)	<.001
Arthritis (%), OR (95% CI)	33.3 (33.3 - 34.3)	66.7 (65.9 - 66.9)	<.001
Current smoking (%), OR (95% CI)	22.2 (21.6 - 22.9)	77.8 (77.0 - 78.7)	<.001
Active alcohol drinking (%), OR (95% CI)	28.6 (28.0 - 29.8)	71.4 (70.8 - 71.9)	<.001
Psychiatric medication (%), OR (95% CI)	35.1 (34.5 - 35.7)	64.9 (64.1 - 65.8)	<.001
Vigorous exercise (%), OR (95% CI)	40.2 (39.4 - 40.9)	59.8 (59.1 - 60.6)	<.001
Household income, mean (95% CI)	62,595.1 (53,582.5 - 70,419.8)	64,289.8 (58,925.3 - 69,643.1)	<.001

[a]OR: odds ratio.

[b]Not applicable.

[c]Other races: Asian, Latino, and Native American.

Ophthalmic Parameters and Falls

Higher rates of any type of falls and serious falls were observed with worse eyesight (see Tables 3 and 4). Approximately half of the participants who reported poor vision or blindness also reported falling, while 42.8% (95% CI 38.3% - 46.7%) of those with blindness had serious injuries associated with the fall. Of the participants with poor distant and near vision, 53.2% (95% CI 52.7% - 54.1%) and 50.2% (95% CI 49.9% - 50.6%) reported falls, respectively, and over a third of the falls were serious. In addition, falls were reported by 39.6% (95% CI 38.8% - 40.5%) of those who stated they had glaucoma.

Table . Ophthalmic parameters of the study sample between patients who reported falls and no falls (total number of participant-observations).

Characteristic	Participants with any self-reported falls (n=40,477), OR[a] (%), (95% CI)	Participants with no self-reported falls (n=77,357), OR (%), (95% CI)	*P* value
Overall	37.9 (37.7 - 40.1)	62.1 (61.9 - 62.3)	—[b]
Overall eyesight			<.001
Excellent	28.7 (28.3-29.7)	71.3 (69.7-71.9)	
Very good	29.1 (28.8-29.5)	70.9 (70.1-71.8)	
Good	32.2 (32.0-32.5)	67.8 (66.9-68.8)	
Fair	38.9 (38.1-40.5)	61.1 (59.7-63.1)	
Poor	51.1 (50.7-51.4)	48.9 (47.6-49.7)	
Blind	48.3 (46.9-49.9)	51.7 (49.1-54.1)	
Distance vision			<.001
Excellent	29.2 (28.1-30.0)	70.8 (69.4-71.9)	
Very good	29.8 (28.8-30.8)	70.2 (69.2-71.2)	
Good	32.8 (31.8-33.8)	67.2 (66.2-67.9)	
Fair	41.8 (41.0-42.9)	58.2 (56.8-59.3)	
Poor	53.2 (52.7-54.1)	46.8 (44.9-47.8)	
Near vision			<.001
Excellent	28.8 (28.3-30.0)	71.2 (69.8-71.9)	
Very good	29.8 (29.1-30.5)	70.2 (69.3-70.9)	
Good	32.2 (32.1-33.0)	67.8 (66.7-68.4)	
Fair	40.2 (39.7-41.0)	59.8 (58.5-60.8)	
Poor	50.2 (49.9-50.6)	49.8 (47.9-51.1)	
Glaucoma			<.001
Yes	39.6 (38.8 - 40.5)	60.4 (58.5 - 62.2)	
No	32.8 (32.0 - 33.8)	67.2 (66.4 - 69.0)	

[a]OR: odds ratio.

[b]Not applicable.

Table . Ophthalmic parameters of the study sample between patients who reported serious falls and no serious falls (total number of participant-observations).

Characteristic	Participants with any self-reported serious falls (n=13,471), OR[a] (%), (95% CI)	Participants with no self-reported serious falls (n=27,006), OR (%), (95% CI)	*P* value
Overall	30.9 (30.7 - 31.1)	69.1 (68.9 - 69.3)	—[b]
Overall eyesight			<.001
Excellent	31.7 (30.1 - 33.7)	68.3 (65.3 - 69.9)	
Very good	30.8 (29.6 - 31.7)	69.2 (67.4 - 70.2)	
Good	31.7 (30.5 - 33.0)	68.3 (67.0 - 69.7)	
Fair	33.3 (32.3 - 34.8)	67.7 (65.1 - 69.1)	
Poor	37.9 (36.8 - 39.6)	62.1 (60.4 - 64.3)	
Blind	42.8 (38.3 - 46.7)	57.2 (51.3 - 63.3)	
Distance vision			<.001
Excellent	30.1 (29.7 - 32.4)	69.9 (67.6 - 71.4)	
Very good	31.7 (30.4 - 33.3)	68.3 (66.7 - 69.5)	
Good	31.8 (29.9 - 32.7)	68.2 (67.3 - 69.2)	
Fair	34.8 (33.5 - 36.2)	65.2 (63.0 - 66.6)	
Poor	39.1 (37.9 - 41.2)	60.9 (58.4 - 62.0)	
Near vision			<.001
Excellent	31.6 (30.5 - 32.6)	68.4 (66.4 - 69.2)	
Very good	31.2 (30.7 - 32.9)	68.8 (67.1 - 69.8)	
Good	32.6 (31.4 - 33.8)	67.4 (67.3 - 68.9)	
Fair	34.2 (33.5 - 35.9)	65.8 (63.1 - 67.6)	
Poor	38.6 (37.6 - 40.6)	61.4 (59.4 - 63.1)	
Glaucoma			<.001
Yes	37.2 (35.4 - 39.5)	62.8 (60.3 - 64.8)	
No	32.3 (32.1 - 33.4)	67.7 (65.7 - 68.9)	

[a]OR: odds ratio.

[b]Not applicable.

Cross-Sectional Analysis

Cross-sectionally, vision impairment was associated with higher odds of falls: impaired overall eyesight (aOR 1.36, 95% CI 1.20 - 1.56; *P*<.001), impaired distance vision (aOR 1.37, 95% CI 1.32 - 1.42; *P*<.001), and impaired near vision (aOR 1.33, 95% CI 1.27 - 1.37; *P*<.001) all conferred significantly increased odds of falls after adjusting for confounding factors (see Table 5 and Figure 1). The odds of serious falls increased in impaired overall eyesight with adjusted ORs of 1.20 (95% CI 1.03 - 1.44; *P*<.001), while having overall eyesight qualifying as legal blindness increased the odds of falls (aOR 1.32, 95% CI 1.26 - 1.37; *P*<.001) and serious falls (aOR 1.16, 95% CI 1.03 - 1.21; *P*<.001) compared to not having impaired vision. Self-reported glaucoma increased the adjusted odds of falls (aOR 1.15, 95% CI 1.07 - 1.24; *P*<.001) and serious falls (aOR 1.15, 95% CI 1.05 - 1.26; *P*<.001).

Table . Cross-sectional model showing the association between falls and ophthalmic parameters.

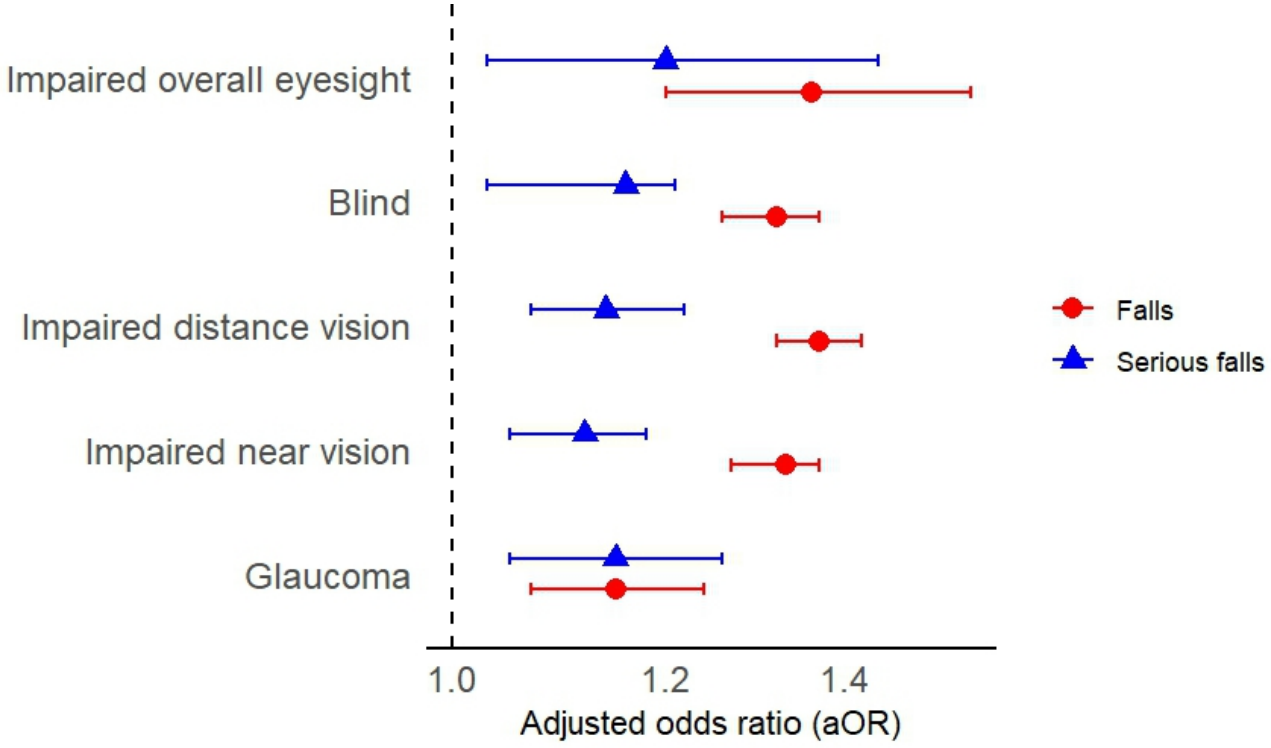
Type of falls	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)[b]	*P* value
Falls				
Overall eyesight				
Not impaired	Ref[c]	Ref	Ref	Ref
Impaired	1.60 (1.52 - 1.73)	<.001	1.36 (1.20 - 1.56)	<.001
Blind	2.03 (1.83 - 2.44)	<.001	1.32 (1.26 - 1.37)	<.001
Impaired distance vision	1.78 (1.70 - 1.89)	<.001	1.37 (1.32 - 1.42)	<.001
Impaired near vision	1.89 (1.69 - 1.99)	<.001	1.33 (1.27 - 1.37)	<.001
Glaucoma	1.33 (1.27 - 1.41)	<.001	1.15 (1.07 - 1.24)	<.001
Serious falls				
Overall eyesight				
Not impaired	Ref	Ref	Ref	Ref
Impaired	1.23 (1.11 - 1.31)	<.001	1.20 (1.03 - 1.44)	.045
Blind	1.49 (1.24 - 1.89)	<.001	1.16 (1.03 - 1.21)	.005
Impaired distance vision	1.22 (1.19 - 1.27)	<.001	1.14 (1.07 - 1.22)	.007
Impaired near vision	1.21 (1.14 - 1.29)	<.001	1.12 (1.05 - 1.18)	<.001
Glaucoma	1.24 (1.12 - 1.33)	<.001	1.15 (1.05 - 1.26)	.004

[a]OR: odds ratio.

[b]Adjusted for age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol, activity, psychiatric medication, vigorous activities, and household income.

[c]Reference.

Figure 1. Forest plot illustrating the association between falls and ophthalmic parameters, based on a cross-sectional multivariable model using generalized estimating equations.



Longitudinal Analysis

In the longitudinal analysis of the association between ophthalmic parameters with falls and serious falls, impaired overall eyesight, impaired distance vision, impaired near vision, and self-reported glaucoma all increased the odds of future falls with an OR of 1.23 (95% CI 1.16 - 1.29; *P*<.001), 1.27 (95% CI 1.20 - 1.38; *P*<.001), 1.23 (95% CI 1.19 - 1.32; *P*<.001), and 1.25 (95% CI 1.13 - 1.37; *P*<.001), respectively (see Table 6 and Figure 2). Impaired overall eyesight, distance vision, and near vision also increased the risk of serious falls in the subsequent 2 years with an OR of 1.18 (95% CI 1.05 - 1.23; *P*<.001), 1.13 (95% CI 1.04 - 1.21; *P*<.001), and 1.15 (95% CI 1.07 - 1.22; *P*<.001), respectively. There was no association between self-reported glaucoma and the odds of future serious falls (*P*=.32).

Table . Longitudinal model showing the association between falls and ophthalmic parameters.

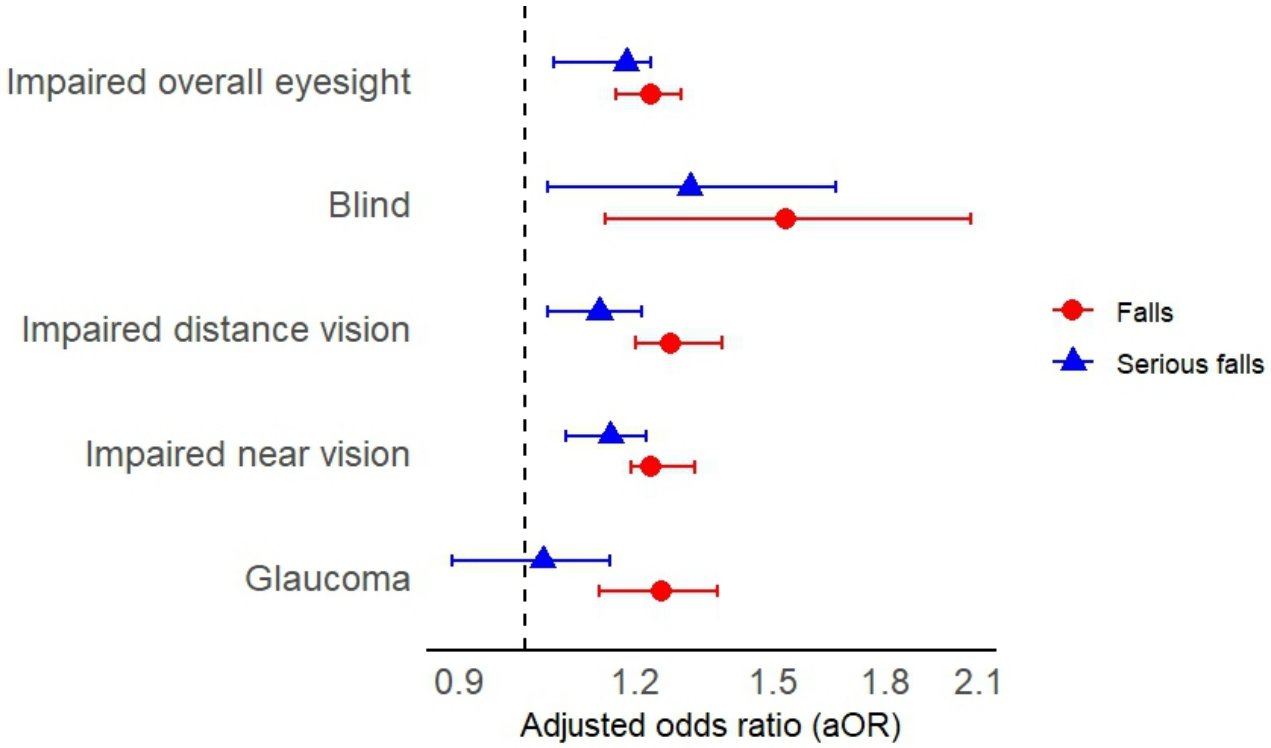
Type of falls	OR[a] (95% CI)	*P* value	Adjusted OR (95% CI)[b]	*P* value
Falls				
Overall eyesight				
Not impaired	Ref[c]	Ref	Ref	Ref
Impaired	1.59 (1.53 - 1.67)	<.001	1.23 (1.16 - 1.29)	<.001
Blind	2.03 (1.65 - 2.68)	<.001	1.53 (1.14 - 2.07)	.005
Impaired distance vision	1.73 (1.63 - 1.88)	<.001	1.27 (1.20 - 1.38)	.001
Impaired near vision	1.69 (1.63 - 1.75)	<.001	1.23 (1.19 - 1.32)	<.001
Glaucoma	1.40 (1.23 - 1.52)	<.001	1.25 (1.13 - 1.37)	<.001
Serious falls				
Overall eyesight				
Not impaired	Ref	Ref	Ref	Ref
Impaired	1.17 (1.10 - 1.23)	<.001	1.18 (1.05 - 1.23)	<.001
Blind	1.45 (1.17 - 1.93)	<.001	1.31 (1.04 - 1.66)	.046
Impaired distance vision	1.20 (1.13 - 1.26)	<.001	1.13 (1.04 - 1.21)	.004
Impaired near vision	1.19 (1.13 - 1.26)	<.001	1.15 (1.07 - 1.22)	<.001
Glaucoma	1.15 (1.13 - 1.17)	.039	1.03 (0.89 - 1.15)	.32

[a]OR: odds ratio.

[b]Adjusted for age, sex, Hispanic ethnicity, race, education, marital status, hypertension, diabetes, heart disease, stroke, arthritis, smoking, alcohol, activity, psychiatric medication, vigorous activities, and household income.

[c]Reference.

Figure 2. Forest plot illustrating the association between falls and ophthalmic parameters, based on a longitudinal multivariable model using Generalized Estimating Equation.



Analysis Between Mean Number of Falls and Eyesight

The mean number of falls was associated with the degree of self-reported vision impairment (see Table 7). The level of eyesight reported in the questionnaire was associated with the reported number of falls in the same survey cycle as well as the subsequent cycle (P<.001 for both). Approximately 2.57 (SD 4.58) falls per person who fell were observed in the group with excellent current eyesight compared to 3.70 (SD 5.33) for those with poor reported vision. Similarly, self-reported excellent overall eyesight had an average of 2.78 (SD 4.17) falls per person while poor overall eyesight had approximately 3.65 (SD 5.40) falls per person in the subsequent cycle.

Table . The association of the degree of visual impairment and mean number of falls.

Degree of impairment	Excellent	Very good	Good	Fair	Poor	*P* value
Cross-sectional, mean (SD)						
Overall eyesight	2.57 (4.58)	2.58 (3.84)	2.74 (4.05)	3.15 (4.59)	3.70 (5.33)	<.001
Distal eyesight	2.66 (4.16)	2.70 (4.04)	2.79 (4.05)	3.34 (4.94)	3.73 (5.26)	<.001
Near eyesight	2.59 (3.86)	2.65 (4.23)	2.77 (4.04)	3.2 (4.52)	3.88 (5.79)	<.001
Longitudinal, mean (SD)						
Overall eyesight	2.78 (4.17)	2.71 (4.10)	2.84 (4.18)	3.19 (4.72)	3.65 (5.40)	<.001
Distal eyesight	2.74 (4.13)	2.75 (3.95)	2.93 (4.40)	3.27 (4.89)	3.59 (5.08)	<.001
Near eyesight	2.66 (3.92)	2.72 (4.01)	2.91 (4.37)	3.28 (4.92)	3.62 (5.10)	<.001

Discussion

Principal Findings

In this large and diverse longitudinal population-based study of older adults, poor vision, including both distance and near vision, increased the odds of falling, and more severe vision loss was related to the number of falls. We additionally found that the presence of self-reported glaucoma was associated with falls.

Comparison With Previous Studies

Many population-based studies have found associations between self-reported vision impairment and falls. The Survey of Health, Ageing, and Retirement in Europe (SHARE) study found that self-reported distance and near vision impairment were cross-sectionally and longitudinally associated with increased risk of falls [21]. The Beaver Dam Eye Study reported that people with worse binocular acuity had higher odds of 2 or more falls [12], while the EPIC-Norfolk eye study found significant associations between poor self-reported distance vision and falls

independent of visual acuity [20]. Similarly, we demonstrate here that the presence of visual impairment, including overall eyesight, near vision, and distance vision, is all associated with significantly increased odds for falls and severe falls in both our cross-sectional and longitudinal analyses. Our results suggest that self-reported visual status is predictive of falls and serious falls.

Importantly, the degree of self-reported visual impairment was associated with the number of falls even after adjusting for potential confounders. When looking at the mean number of falls that occur per person for each level of reported vision, we found that though people with excellent vision had falls, the number increased significantly as vision worsened, with those with blindness having the highest risk of falls. This is true for both falls that occurred 2 years before and those that occurred in the following 2 years of the reported vision and does not take into account the fact that blindness likely reduces the number of steps taken.

This study also subclassified falls as serious based on whether a fall had any associated injuries requiring medical treatment. The incidence of serious falls increased with age and deterioration of eyesight, and we found the risk factors for falls and serious falls to be similar. Other studies, including a Finnish population-based study, have examined risk factors of major injurious falls and similarly found poor distant visual acuity to be a risk factor for major falls among disabled elder adults [22].

Visual input plays a critical role in coordinating and planning movement, as well as maintaining balance [23]. Individuals with poor vision are more likely to lose stability, alter their gait to avoid obstacles, and ultimately increase their risk of slips, trips, and falls [24]. Our study differs from previous research in this area, which has largely relied on data from previous falls and visual function assessments. By using self-reported data, our research offers a fresh perspective and highlights the potential of self-reported visual function data as a predictive tool for falls risk.

We also found that glaucoma is a risk factor for serious falls resulting in injury despite the fact that our definition of glaucoma was imperfect (based on self-report only). Any nondifferential misclassification in diagnosing glaucoma would have driven our associations toward the null, and this implies that the associations found were real and may have been underestimated. Previous studies have also reported glaucoma as an independent risk factor for falling [13]. Glaucoma severity has been reported as a risk factor for falls, with advanced glaucoma patients having the highest risk of falls compared to controls with no glaucoma and mild or moderate glaucoma [25]. While central vision and, therefore, visual acuity are often preserved in glaucoma until severe stages of disease, glaucoma patients experience changes in multiple other aspects of vision, including VF loss, decrease in contrast sensitivity, and loss of depth perception. All these visual measures have been shown to independently increase the risk of falling [13,14,26]. Indeed, the Salisbury Eye Evaluation found that loss in peripheral VF is a more important risk factor for falls than loss of central VF [14]. Patients with glaucoma may experience one or all visual deficits simultaneously, further increasing the risk of falling.

Studies have suggested various interventions to reduce the risk of falls, including mobility training, vision interventions to improve vision (eye examinations, cataract surgery, and prescription glasses) [27], risk factor assessment, and environmental modifications to make living environment less hazardous [28]. Visual impairment leads to physical impairment of mobility and balance control, increasing the risk of falls [29]. Glaucoma initially affects peripheral vision through retinal nerve fiber layer loss, which is associated with decreased postural stability; therefore, glaucoma patients may benefit from mobility training to reduce falling risks [30,31].

Clinical Implications

Our results provide evidence to support the development of multifaceted fall prevention strategies. Specifically, our findings suggest that self-reported vision screening could be a valuable tool in identifying individuals at risk of falls. This screening could be easily integrated into telemedicine platforms or incorporated into standard clinical guidelines. Furthermore, we propose the development of a simple, score-based questionnaire that can be used to predict falls risk in older adults. Those with higher scores could be targeted for earlier intervention, thereby reducing their risk of falls.

Limitations

There were several limitations to our study. All data used in this study were from self-reported questionnaires administered every 2 years and are susceptible to recall bias and recall error. Specifically, participants may have difficulty recalling their visual function or falls history, which could lead to inaccurate reporting. In addition, while self-reported visual function has been shown to correlate closely with objective visual function, it is also understood as a multidimensional measure that reflects some, but not all components of vision [19]. Since no eye examinations were done and objective visual function was not measured, we cannot determine which visual components contributed to participants' self-reported visual function. This study is also limited by the specific questions collected as part of the HRS study. No questions were asked about the other leading causes of blindness and visual impairment in this older adult population, such as age-related macular degeneration and other neurological conditions that could affect falls, such as Parkinson disease, dementia, or epilepsy. Finally, we could not explore other known vision-related risk factors for falls such as contrast sensitivity, VF deterioration, depth perception, and stereoacuity [26]. Nonvision factors such as cognitive ability, balance, motor control, and degree of mobility, which are known to be associated with falls, were not evaluated since they were not included in the questionnaire. The omission of these factors may limit the generalizability of our findings and highlights the need for future studies to incorporate a more comprehensive set of measures.

Conclusions

Our analysis of data from a large, longitudinal national survey of older adults demonstrated that vision-related risk factors, particularly the presence and degree of visual impairment and glaucoma, increase the rate of falls and serious falls. We additionally show that self-reported vision status correlates with

the number of fall events both at the time of and occurring after the survey date. Our results support screening for the presence of visual impairment and signs of glaucoma in older adults to identify individuals at the highest risk of falling who might most benefit from fall prevention interventions.

Acknowledgments

The Health and Retirement Study is supported by the National Institute on Aging, supplemented by the Social Security Agency, and operated from the Institute for Social Research at the University of Michigan.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Self-reported visual function and falls questions in the Health and Retirement Study (HRS).
[DOCX File, 15 KB - aging_v8i1e68771_app1.docx]

References

1. Step safely: strategies for preventing and managing falls across the life-course. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/978924002191-4 [accessed 2025-07-02]
2. Bergen G, Stevens MR, Burns ER. Falls and fall injuries among adults aged ≥65 years - United States, 2014. MMWR Morb Mortal Wkly Rep 2016 Sep 23;65(37):993-998. [doi: 10.15585/mmwr.mm6537a2] [Medline: 27656914]
3. Moreland B, Kakara R, Henry A. Trends in nonfatal falls and fall-related injuries among adults aged ≥65 years - United States, 2012-2018. MMWR Morb Mortal Wkly Rep 2020 Jul 10;69(27):875-881. [doi: 10.15585/mmwr.mm6927a5] [Medline: 32644982]
4. Rubenstein LZ, Josephson KR. The epidemiology of falls and syncope. Clin Geriatr Med 2002 May;18(2):141-158. [doi: 10.1016/s0749-0690(02)00002-2] [Medline: 12180240]
5. Tinetti ME. Clinical practice. Preventing falls in elderly persons. N Engl J Med 2003 Jan 2;348(1):42-49. [doi: 10.1056/NEJMcp020719] [Medline: 12510042]
6. Facts about falls. : CDC; 2023 URL: https://www.cdc.gov/falls/data-research/facts-stats/index.html [accessed 2025-07-02]
7. Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016 Oct;388(10053):1459-1544. [doi: 10.1016/S0140-6736(16)31012-1]
8. Close J, Ellis M, Hooper R, Glucksman E, Jackson S, Swift C. Prevention of falls in the elderly trial (PROFET): a randomised controlled trial. Lancet 1999 Jan 9;353(9147):93-97. [doi: 10.1016/S0140-6736(98)06119-4] [Medline: 10023893]
9. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]
10. Dhargave P, Sendhilkumar R. Prevalence of risk factors for falls among elderly people living in long-term care homes. Journal of Clinical Gerontology and Geriatrics 2016 Sep;7(3):99-103. [doi: 10.1016/j.jcgg.2016.03.004]
11. Patil S, Suryanarayana S, Shivraj N, Murthy N, Dinesh R. Risk factors for falls among elderly: a community-based study. Int J Health Allied Sci 2015;4(3):135. [doi: 10.4103/2278-344X.160867]
12. Klein BEK, Moss SE, Klein R, Lee KE, Cruickshanks KJ. Associations of visual function with physical outcomes and limitations 5 years later in an older population: the Beaver Dam eye study. Ophthalmology 2003 Apr;110(4):644-650. [doi: 10.1016/S0161-6420(02)01935-8] [Medline: 12689880]
13. Ivers RQ, Cumming RG, Mitchell P, Attebo K. Visual impairment and falls in older adults: the Blue Mountains Eye Study. J Am Geriatr Soc 1998 Jan;46(1):58-64. [doi: 10.1111/j.1532-5415.1998.tb01014.x] [Medline: 9434666]
14. Freeman EE, Muñoz B, Rubin G, West SK. Visual field loss increases the risk of falls in older adults: the Salisbury eye evaluation. Invest Ophthalmol Vis Sci 2007 Oct;48(10):4445-4450. [doi: 10.1167/iovs.07-0326] [Medline: 17898264]
15. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. JAMA 1989 May 12;261(18):2663-2668. [Medline: 2709546]
16. Campbell AJ, Borrie MJ, Spears GF. Risk factors for falls in a community-based prospective study of people 70 years and older. J Gerontol 1989 Jul;44(4):M112-M117. [doi: 10.1093/geronj/44.4.m112] [Medline: 2738307]
17. Lord SR, Ward JA, Williams P, Anstey KJ. Physiological factors associated with falls in older community‐dwelling women. J American Geriatrics Society 1994 Oct;42(10):1110-1117. [doi: 10.1111/j.1532-5415.1994.tb06218.x]

18. Fisher GG, Ryan LH. Overview of the health and retirement study and introduction to the special issue. Work Aging Retire 2018 Jan;4(1):1-9. [doi: 10.1093/workar/wax032] [Medline: 29423243]
19. El-Gasim M, Munoz B, West SK, Scott AW. Associations between self-rated vision score, vision tests, and self-reported visual function in the Salisbury Eye Evaluation Study. Invest Ophthalmol Vis Sci 2013 Sep 27;54(9):6439-6445. [doi: 10.1167/iovs.12-11461] [Medline: 23812494]
20. Yip JLY, Khawaja AP, Broadway D, et al. Visual acuity, self-reported vision and falls in the EPIC-Norfolk Eye study. Br J Ophthalmol 2014 Mar;98(3):377-382. [doi: 10.1136/bjophthalmol-2013-304179] [Medline: 24338086]
21. Ogliari G, Ryg J, Qureshi N, Andersen-Ranberg K, Scheel-Hincke LL, Masud T. Subjective vision and hearing impairment and falls among community-dwelling adults: a prospective study in the Survey of Health, Ageing and Retirement in Europe (SHARE). Eur Geriatr Med 2021 Oct;12(5):1031-1043. [doi: 10.1007/s41999-021-00505-4] [Medline: 34003480]
22. Koski K, Luukinen H, Laippala P, Kivelä SL. Risk factors for major injurious falls among the home-dwelling elderly by functional abilities. A prospective population-based study. Gerontology 1998;44(4):232-238. [doi: 10.1159/000022017] [Medline: 9657085]
23. Manchester D, Woollacott M, Zederbauer-Hylton N, Marin O. Visual, vestibular and somatosensory contributions to balance control in the older adult. J Gerontol 1989 Jul;44(4):M118-M127. [doi: 10.1093/geronj/44.4.m118] [Medline: 2786896]
24. Buckley JG, Panesar GK, MacLellan MJ, Pacey IE, Barrett BT. Changes to control of adaptive gait in individuals with long-standing reduced stereoacuity. Invest Ophthalmol Vis Sci 2010 May;51(5):2487-2495. [doi: 10.1167/iovs.09-3858] [Medline: 20335609]
25. Bhorade AM, Perlmutter MS, Sabapathypillai SL, Goel M, Wilson B, Gordon MO. Rate of falls, fear of falling, and avoidance of activities at-risk for falls in older adults with glaucoma. Am J Ophthalmol 2021 Jul;227:275-283. [doi: 10.1016/j.ajo.2021.02.017] [Medline: 33626364]
26. Mehta J, Czanner G, Harding S, Newsham D, Robinson J. Visual risk factors for falls in older adults: a case-control study. BMC Geriatr 2022 Feb 17;22(1):134. [doi: 10.1186/s12877-022-02784-3] [Medline: 35177024]
27. Zhang XY, Shuai J, Li LP. Vision and relevant risk factor interventions for preventing falls among older people: a network meta-analysis. Sci Rep 2015 May 28;5(1):10559. [doi: 10.1038/srep10559] [Medline: 26020415]
28. Sotimehin AE, Yonge AV, Mihailovic A, et al. Locations, circumstances, and outcomes of falls in patients with glaucoma. Am J Ophthalmol 2018 Aug;192:131-141. [doi: 10.1016/j.ajo.2018.04.024] [Medline: 29750950]
29. Zwierko T, Jedziniak W, Florkiewicz B, et al. The consequences of glaucoma on mobility and balance control in the older adults: a cross-sectional study. J Aging Phys Act 2021 Jun 1;29(3):372-381. [doi: 10.1123/japa.2020-0079] [Medline: 32994380]
30. O'Connell C, Mahboobin A, Drexler S, et al. Effects of acute peripheral/central visual field loss on standing balance. Exp Brain Res 2017 Nov;235(11):3261-3270. [doi: 10.1007/s00221-017-5045-x] [Medline: 28765993]
31. Black AA, Wood JM, Lovie-Kitchin JE, Newman BM. Visual impairment and postural sway among older adults with glaucoma. Optom Vis Sci 2008 Jun;85(6):489-497. [doi: 10.1097/OPX.0b013e31817882db] [Medline: 18521027]

Abbreviations

aOR: adjusted odds ratio
HRS: Health and Retirement Study
OR: odds ratio
VF: visual field

Edited by R Bjarnadottir; submitted 24.11.24; peer-reviewed by P Roberts, Q Mao, Y Gu; revised version received 11.03.25; accepted 21.04.25; published 31.07.25.

Please cite as:
Seresirikachorn K, Somyanonthanakul R, Johnson M, Singhanetr P, Gatedee J, Friedman D, Zebardast N
The Impact of Vision Impairment on Self-Reported Falls Among Older US Adults: Cross-Sectional and Longitudinal Study
JMIR Aging 2025;8:e68771
URL: https://aging.jmir.org/2025/1/e68771
doi:10.2196/68771

Original Paper

Building Strong Foundations: Nonrandomized Interventional Study of a Novel, Digitally Delivered Fall Prevention Program for Older Adults

David Wing[1,2], MS; Jeanne F Nichols[1,2], PhD; Hava Shoshana Barkai[1,2], MS; Olivia Culbert[1,2], MS; Daniel Moreno[1,2], MS; Michael Higgins[1,2], MS; Anna O'Brien[1,2], BS; Mariana Perez[2], BS; Hope Davey[2], BS; Ryan Moran[2,3], MPH, MD

[1]Herbert Wertheim School of Public Health and Human Longevity, University of California, San Diego, San Diego, CA, United States

[2]Exercise and Physical Activity Resource Center, University of California, San Diego, San Diego, CA, United States

[3]School of Medicine, University of California, San Diego, San Diego, CA, United States

Corresponding Author:
David Wing, MS
Exercise and Physical Activity Resource Center
University of California, San Diego
9500 Gilman Drive #0811
CALIT2/Atkinson Hall Room 3504
San Diego, CA, 92093
United States
Phone: 1 858 534 9315
Fax: 1 858 534 9404
Email: dwing@eng.ucsd.edu

Abstract

Background: Injuries from falls are a major concern among older adults. Targeted exercise has been shown to improve fall risk, and recommendations for identifying and referring older adults for exercise-based interventions exist. However, even when very inexpensive or free, many do not use available fall prevention programs, citing barriers related to convenience and safety. These issues are even greater among older adults residing in rural areas where facilities are less abundant. These realities highlight the need for different approaches to reducing falls in novel ways that increase reach and are safe and effective. Web-based delivery of exercise interventions offers some exciting and enticing prospects.

Objective: Our objective was to assess the efficacy of the Strong Foundations exercise program to change markers of physical function, posture, balance, strength, and fall risk.

Methods: Strong Foundations is a once weekly (60 minutes), 12-week iterative program with 3 core components: postural alignment and control, balance and mobility, and muscular strength and power. We used a quasi-experimental design to determine changes in physical function specific to balance, postural control, and muscular strength among older adults at low or moderate risk of falling.

Results: A total of 55 low-risk and 37 moderate-risk participants were recruited. Participants significantly improved on the 30-second Chair Stand (mean change of 1, SD 3.3 repetitions; P=.006) and Timed Up and Go (mean change of 0.2, SD 0.7 seconds; P=.004), with the moderate-risk group generally improving to a greater degree than the low-risk group. Additionally, Short Physical Performance Battery performance improved significantly in the moderate-risk category (P=.02). The majority of postural measures showed statistically significant improvement for both groups (P<.05). Measures of “relaxed” posture showed improvements between 6% and 27%. When an “as tall as possible” posture was adopted, improvements were ~36%.

Conclusions: In this 12-week, iterative, web-based program, we found older adults experienced improvement not only in measures used in clinical contexts, such as the 30-second Chair Stand and Timed Up and Go, but also contextualized gains by providing deeper phenotypical measurement related to posture, strength, and balance. Further, many of the physical improvements were attenuated by baseline fall risk level, with those with the highest level of risk having the greater gains, and, thus, the most benefit from such interventions.

(JMIR Aging 2025;8:e68957) doi:10.2196/68957

KEYWORDS

exercise; older adults; digital intervention; Zoom; balance; posture; strength; fall prevention

Introduction

The absolute number of adults older than 65 years of age has been steadily growing for decades and now outnumbers the number of children younger than 18 years of age [1]. In addition, estimates indicate that older adults will represent greater than 23% of the population in the United States by 2035 [1]. Injuries resulting from falls are the largest cause of accidental death and mobility-related disability among older adults. This is likely because nearly one-quarter of community-residing adults older than 65 years of age fall annually, and that number rises to almost half of those older than 80 years of age [2]. Particularly concerning is the fact that once an individual starts to fall, there are often additional falls. Indeed, evidence from numerous sources suggests that between 10% and 44% of older adult patients who have fallen will sustain additional falls in the following year [3-7].

Targeted strength and balance exercises have consistently been shown to improve fall risk, and accordingly, recommendations for identifying and referring older adults for exercise-based interventions have been developed. Public health and clinically oriented authorities, including the Centers for Disease Control and Prevention (CDC) and the US Preventive Services Task Force, have acknowledged the importance of risk assessment and physical activity (PA) in improving fall risk for older adults. The CDC's Stopping Elderly Accidents and Deaths Initiative (STEADI) has also established an evidenced-based questionnaire that assesses known risk factors to help clinicians stratify individuals by risk category to identify those who should be best served by fall prevention programs [8-10].

Many exercise programs for older adults exist. While some health insurance may include access to programs such as Silver Sneakers and EnhanceFitness PA at no additional charge [11], many older adults may have to pay substantial fees to participate. Further, even when very inexpensive or free, many eligible older adults do not use these programs, citing several barriers [11,12]. These include personal preferences (dislike of gymnasiums and lack of knowledge regarding appropriate activity), environmental factors (difficulty in reaching gymnasiums or classes, concerns about bad weather and driving, concerns about being "old" in a gymnasium environment), and structural factors (limited number of facilities and concerns about instructor expertise with older adults) [12]. These issues are even greater among older adults residing in rural areas. Rural residents have fewer facilities for guided PA, more limited transportation infrastructure [13], and less exercise expertise than their urban counterparts [14]. Confounding these difficulties, ongoing concerns exist surrounding the COVID-19 transmissibility alongside the availability, and desirability, of in-person programs. Since the height of the pandemic, there has been a growth in available options for in-person exercise programs. However, while vaccines and exposure have decreased the likelihood of serious infection, there remains some risk associated with (large) in-person gatherings, particularly among the more vulnerable older adult population.

These realities highlight the need for different approaches in addressing the public health challenge of reducing the likelihood of falls in older adults in novel ways that increase reach and are safe and effective. Digital delivery of exercise interventions offers some exciting and enticing prospects, as not only can this platform eliminate barriers related to difficulty with transportation and concerns about the gymnasium environment, but it can also bring highly qualified experts to individuals who are the most in need. These include those experiencing specific conditions that require special instruction and those living in rural and geographically remote areas. Given that older adults are increasingly confident and competent using the internet and videoconferencing tools [15,16], developing interventions that leverage this technology to improve access to fall-risk reduction exercise programs, including balance, posture, and strength training, is both worthwhile and feasible.

Despite these possibilities, there remains limited research on digital group–based exercise programs, particularly those focused on fall prevention. Further, the available evidence is generally drawn from relatively small samples [17-19], or the studies have been conducted in populations with specific medical conditions or comorbidities such as diabetes [20,21] or mild cognitive impairment [22]. To the authors' knowledge, there are only 2 randomized studies with a sample size greater than 20 that have focused on healthy older adults. Although the exercise modalities differed, following the intervention, participants did have generally positive findings regarding improvements in strength and balance [23,24]. However, neither intervention included exercises specific to postural control, a factor strongly hypothesized to influence fall risk [25,26]. Moreover, instruction during exercise sessions in nearly all studies we reviewed, regardless of the health status of the participants, was generalized to the entire class due to the difficulty of the instructor speaking to any 1 person to correct technique without interrupting the flow of the class.

Therefore, we developed and digitally deployed an evidence-based exercise program, Strong Foundations (SF), with instruction provided in real time, to enhance physical function and ultimately prevent falls among older adults. The aim of this study was to evaluate the SF program's effectiveness in improving the posture, balance, and strength of older adults using valid, reliable, and widely used measures of physical function.

Methods

Ethical Considerations

This study was approved by the institutional review board (#806696) at the University of California, San Diego, and informed consent was gathered from all participants. They were further informed that they did not have to complete any measurements and were free to withdraw from the exercise sessions at any time. Participants were additionally provided with US $20 for each laboratory visit to cover travel costs and were given resistance bands to be used during the exercise

intervention (value ~US $10). All data were de-identified prior to analyses with a code to participant numbers and identities held only by the PI and study coordinator and maintained in protected file servers.

Study Design and Recruitment

We used a quasi-experimental design to determine the efficacy of digitally delivered exercise for improving physical function specific to balance, postural control, and muscular strength among older adults at low or moderate risk of falling. Several strategies were used to recruit participants: (1) posters displayed at older adult living facilities and community centers; (2) emails sent to existing lists of older adults who previously expressed interest in research studies in general, and in fracture prevention, specifically; and (3) announcements made following educational presentations on fall and fracture prevention delivered digitally or in person to various groups of older adults in the community. Those who expressed interest were contacted by study personnel and screened by phone for eligibility. Inclusion criteria included being 60 years and older, having an internet-enabled device with a screen of 7+ inches (tablet, laptop, or similar), and scoring less than 8 on the CDC's STEADI questionnaire. Participants were further stratified based on their answers to the STEADI questionnaire into low- and moderate-risk groups with a maximal score of 3 and 7, respectively [27].

Intervention Design

Designed by 2 exercise physiologists working in consultation with both a medical doctor and a doctor of physical therapy, the SF program is a once weekly (60 minutes), 12-week iterative program with 3 core components: postural alignment and control, balance and mobility, and muscular strength and power. All instructors leading the training had, at minimum, a bachelor's degree in an associated field (kinesiology, exercise physiology, etc) and licensing as a personal trainer from an accredited national institution (American College of Sports Medicine, American Council on Exercise, and National Academy of Sports Medicine). In addition to the weekly group class, participants received both printed and recorded instruction regarding how to safely complete the exercises without supervision and were "assigned" homework materials and encouraged to practice 2 or more additional times per week preferably with a family member or friend present for safety.

All of the exercises introduced throughout the course were designed to be appropriate for an older adult population and were standardized so that participants received the same basic instruction, but the level of difficulty was scaled (and coached) based on individual capability, experience, and musculoskeletal limitations. Each week, new foundational exercises identified as being important to maintaining strength, postural control, balance, and mobility, and also relevant to specific daily activities (eg, picking up an object from the floor to place it on a countertop) were added to the class. In addition, over time, most exercises initially taught as isolated movements were expanded to more complicated, multijoint and multiplanar compound movements, once again with the intent to mimic daily movement.

Home Setup and Safety Concerns

One to two weeks before the first exercise class, participants were emailed instructions for setting up their exercise space at their residence. This included having, at minimum, a 6×6 ft uncluttered exercise space on either a nonslippery floor or an area with wall-to-wall carpet (no loose rugs) and adequate open space surrounding the designated exercise space. In addition, if possible, they were encouraged to have a small amount of wall space immediately behind their exercise area to allow for intermittent wall exercises. Participants were also provided with a set of resistance bands with the level of difficulty determined based on their self-described strength and were instructed to ensure that they had a chair without wheels (and preferably without arms) available for chair-based exercises and to assist with balance during standing exercise.

Intervention Delivery

One of the novel features of this program is the delivery of semi-individualized instruction in real time within a small group setting. To this end, classes were designed to have 10-15 participants. A single lead instructor provided verbal instruction while demonstrating each exercise. Simultaneously, at least 1, and sometimes 2, additional instructors provided individualized instruction to participants who were performing an exercise incorrectly or those who were ready to move on to a more difficult progression, with a particular emphasis on exercise form and safety. This allowed all participants to develop competency with key exercises while allowing those with experience and confidence to progress to more advanced movements in a timely manner.

In-person classes were offered once per week and lasted slightly longer than 1 hour. Classes began with approximately 5 minutes of aerobic warm-up during, which participants were encouraged to move somewhat vigorously in order to raise their heart rate and elicit blood flow to the working muscles. Following the warm-up, each class had different foci as shown in Table 1. Regardless of the focus of the week, 5-10 minutes were spent teaching the week's foundational exercises with particular emphasis on explaining proper form. In the early weeks of the intervention, participants were instructed to complete the exercises slowly, paying particular attention to ensure that they were feeling the target muscles activate or engage in the appropriate anatomical area. Individualized feedback from an assistant instructor correcting errors in technique was emphasized during this time, making sure that each participant received feedback from an instructor at least once during the class. As the weeks progressed, the lead instructor cued participants to increase the resistance or increase the speed, or increase both the speed and resistance concurrently during strength training exercises, as tolerated, and also increase balance challenges if they felt safe doing so.

Table 1. Weekly foci and foundational exercises.

Week	Foundational exercise 1	Foundational exercise 2	Foundational exercise 3	Foundational exercise 4	Focus of the week
1	Neutral spine or marionette pose	Pelvic tucks or tilts	Hip hikes	Hip hinge	Posture
2	Quick feet drills	Review hip hinge	"Traditional abs"	Self-test: static balance	Balance
3	Front and lateral arm raises (with band)	Hip hinge (review) with squat or chair sit or stand	Reverse lunge	Heel toe raises	Strength
4	Neutral spine and pelvic tuck or tilt—standing	Static balance	Reverse lunge with progression	Heel toe raises with tennis ball	Combine 3 pillars
5	Sit and stand transitions	Around the clock steps	Hip hinge with choice of W/T/Y/I arm extensions	Modified or regular jumping jacks	Floor to stand transition—floor exercises
6	Hip hinge (review) with squat or chair sit or stand	Lunge with chair progression	Single leg heel raises	Hing hinge+picking up objects	Compound strength movements
7	Drinking bird	Squat to lateral leg raise	Single leg heel raises	Toe raise walk around chair	Increase balance challenge
8	Warrior 2 with chair sit	Wall sit with wood chop	Introduce pivot	Quick feet multidirectional	Odd impact or multidirectional movements
9	Weight transfer with split stance	Squat into a high knee and arm reach	Pivot	Head rotations in a tandem stance	Combine balance with compound movements
10	Single leg stance with band pull downs and serial 7's	Lunge with chair support and head rotation	Standing superman with cognitive challenge	Goal post with arm slides	Introduce cognitive challenge
11	Reaching squats	Lateral jacks	Vertical push-ups	Knee drivers	Speed or power movements
12	N/A[a]	N/A	N/A	N/A	Collaborative workout

[a]N/A: not applicable.

For the remainder of the class, a variety of exercises focused on strength, balance, and posture (alone and in combination) in line with the focus of the week were completed in sets of 3 to 5 with pauses between sets to cue the next set of exercises. During the final 5 minutes, participants were guided through a brief cool down that incorporated stretching of the muscles used (most) during the exercise session.

Participants were also provided a secure link to a public-facing web-based video library that hosted a recording of the exercise session and a written description of the foundational exercises emphasized each week. Participants were further encouraged to complete 2 additional exercise sessions weekly. They also were provided a "mid-week challenge" 3 days after the weekly class session. These challenges included video presentations on the importance of posture in preventing falls, selected yoga-based movements for balance, and other short (5-10 minutes) presentations designed to keep participants engaged with the larger fall prevention focus of SF. Finally, participants were reminded of the importance of cardiovascular exercise and encouraged to walk or do other aerobic activities regularly.

Laboratory Measures

Overview

All measures were completed at baseline, prior to the first exercise class. Follow-up visits were completed anytime from the day after the 11th class through 2 weeks following the completion of the last (12th) exercise session. Participants completed their physical measures in the following order:

Balance and Physical Function

In-laboratory testing of balance and physical function included the Short Physical Performance Battery (SPPB) [28] combined with computerized dynamic posturography to quantify postural sway and the Timed Up and Go (TUG) [29].

Participants completed the standard SPPB, which includes (1) a standing balance challenge, in which the participant is asked to stand as still as possible with an increasingly narrow base of support; (2) regular walking speed at a "regular" pace on a 4-m course; and (3) leg strength gathered from the time taken to complete 5 chair stands. Continuous scores were normalized with a maximal score of 12 (4 points per measurement category) [28]. In an effort to address possible ceiling effects associated with the balance component of the SPPB, the BTrackS Balance Plate (version 7.5; Balance Tracking Systems) was used to capture overall postural sway, measured in centimeters.

Participants completed the TUG at both a normal walking speed and as fast as possible. Participants began in a seated position and were instructed to rise without using their arms to push off, walk 3 m, negotiate around an obstacle (rubber cone), return to their chair, and sit back down. Participants were allowed 1 practice trial and then completed each test twice with the best time (ie, lowest) score used for statistical analysis.

Postural Assessments

Posture was measured using a variety of methods designed to identify areas of postural deficiency. Specifically, for general posture, standing height was measured in two ways using a stadiometer (SECA 213) while participants were instructed to (1) "stand as you normally do" and (2) "stand as tall as possible." Both measurements were taken at the top of the breath cycle. To identify cervical and upper thoracic postural decline, the occiput to wall distance (OWD) [30] was measured in the same 2 manners (ie, regular and tall stance). More specifically, OWD was measured by having participants stand against a wall with their heels and buttocks touching the wall or as close as they could get comfortably. They were then instructed to look straight ahead (neutral head position) and hold still while the distance from the occiput to the wall was measured in centimeters. Finally, the Debrunner kyphometer and flexiruler were used to measure thoracic and lumbar curvature using previously published methodology [31,32]. Briefly, the Debrunner kyphometer provided an angle of curvature of the thoracic spine when placed in the joint space between T2 and T3 on one end and T12 and L1 on the other while participants stood in a normal or relaxed manner. The flexiruler is moldable and is placed with one end at the base of C7 and the other in the joint space between L5 and S1. The molded ruler is then traced onto a piece of paper, from which the kyphotic index was calculated by 2 independent raters as a function of the thoracic length and thoracic width using standardized procedures previously described (width/length×100) [31]. The average of the 2 raters was used for analyses.

Functional Muscular Strength

All participants completed a 30-second Chair Stand (30CS) and grip strength test. A subset (initial 3 cohorts of recruited participants) also completed isometric strength testing of hamstrings or quadriceps and trunk or lumbar muscle groups (procedure described in Isometric Muscular Strength section).

The 30CS used a chair with a seat height of 17 inches. Participants were instructed to keep their feet flat on the floor with their arms folded across their chest and touching their chest throughout the test. To be counted, the participant had to rise to a fully upright position and then return to a seated position while maintaining a fairly vertical body position (avoiding excessive forward lean). If participants did not maintain proper form, they were coached, but time was not stopped.

Hand Grip Strength

Hand grip strength was measured using an adjustable grip strength dynamometer (Jamar Plus Digital Hand Dynamometer). During the baseline visit, the grip or handlebar of the dynamometer was adjusted so the second joint of the fingers fits around the handle with handle size (1-5) recorded and used during subsequent visits. Participants were familiarized with the measurement by performing one submaximal effort on each hand. After becoming comfortable with the procedure, participants were instructed to hold the dynamometer with their arms at their side. They were then coached to take a deep breath in and squeeze as hard as possible as they exhaled. Measurement staff provided encouragement throughout each attempt. The measurement was repeated twice on each hand, alternating between the dominant and nondominant hand, with the highest score for each hand recorded in kilograms.

Isometric Muscular Strength

Isometric strength of the hamstrings, quadriceps, and spinal extensor muscles was measured using the Biodex System 4 PRO dynamometer (version 4.60; Biodex Medical Systems). Specifically, isometric strength of the hamstring and quadriceps muscle groups was measured with the leg held in at 45, 75, and 90 degrees of knee flexion. Participants performed 3 repetitions of 5-second maximal isometric contraction, extension followed by flexion, with 5 seconds of rest between repetitions at each angle. A 2-minute rest period was given between sets or leg positions. To decrease participant burden, only 1 leg was tested with participants indicating if either of their legs or knees had any previous injuries, or any current pain, after which the "healthiest" underwent testing. The average maximal contraction at each angle was recorded, and scores were combined into a flexion and extension composite score.

The Biodex Dual Position Back Ex/Flex Attachment, seated-compressed variation (isolated lumbar position), was used to measure the isometric strength of the trunk at 0 degrees of spinal flexion-extension (seated up-right). One set of 3 repetitions of 5-second maximal isometric contractions for both trunk flexion and extension was completed by participants. A 10-second rest was given between repetitions. To minimize the risk of injury, participants were instructed to slowly and gradually increase their muscle engagement to a volitional maximum and not attempt any sudden or explosive movement against the resistance.

Aerobic Fitness

Participants were asked to walk continuously for 2.5 minutes at their normal walking speed on a 50-m corridor [33]. During the walk, participants wore a chest strap–based heart rate monitor (Polar H10). Total distance, average heart rate, and peak heart rate were recorded.

Statistical Analyses

All statistical analyses were conducted using SPSS (version 27; IBM Corp). As this was largely designed as a feasibility study to determine the acceptability of this novel intervention delivery method and the degree to which the intervention could improve metrics of strength, balance, and posture, power calculations were not conducted. Instead, the sample size was primarily determined by safety concerns involving class size, which we settled on 10-15 per class after conducting 3 brief, 4-week pilot sessions beginning with ~6 participants, increasing to 8-12 to keep class size manageable. The sample size was also restricted by fiscal constraints.

Descriptive statistics (percentages, means, and SDs) were used to characterize demographic variables and identify potential outliers. Change scores were derived by subtracting baseline values from follow-up values on an individual level.

Independent 2-tailed *t* tests were conducted to assess differences across groups at baseline. Paired sample *t* tests were conducted to evaluate the effects of the intervention on the measures of

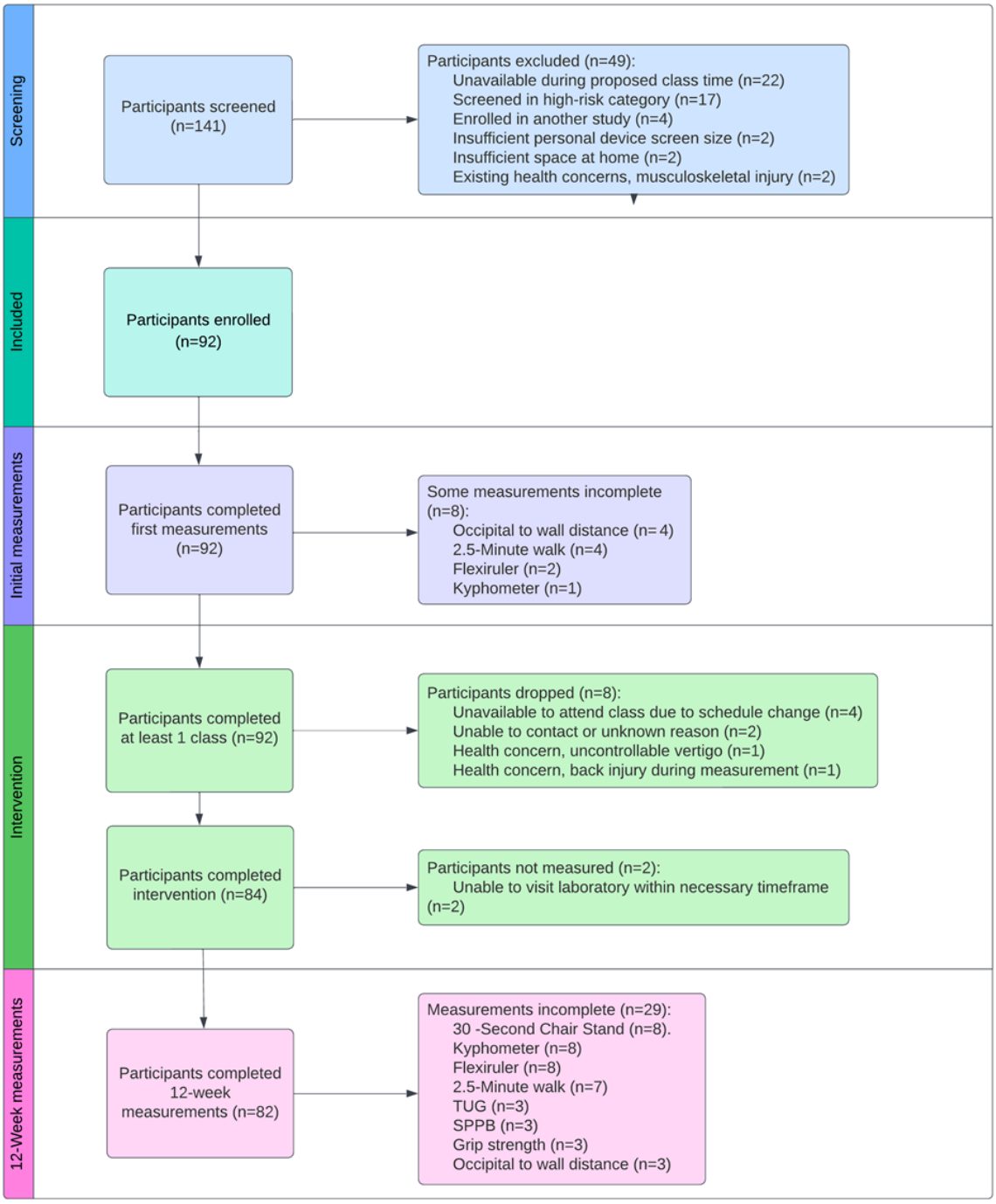
interest, with analyses conducted both for the total population and also with participants divided into fall-risk groups, which were assessed independently of each other. Associations between changes in measures of strength, balance, and posture and number of classes attended were examined using Pearson correlation without controlling for any covariates. When associations were observed, univariate linear modeling was conducted, with sex, chronological age, and baseline score included as covariates. These were included based on the known systematic declines in performance in older individuals, the possibility of systematic differences between male and female individuals, and the differences in potential change associated with baseline performance.

Results

Overview

The number of participants screened, enrolled, and followed through the final measurement is shown in Figure 1.

Figure 1. Participant recruitment flow. The sum of the breakdown of excluded participants at the initial and 12-week measurements exceeds the total number of excluded participants, as some individuals missed more than one assessment. SPPB: Short Physical Performance Battery; TUG: Timed Up and Go.



Population Descriptive Data

A total of 92 participants were enrolled across 8 cohorts with 55 low-risk and 37 moderate-risk individuals assigned to their own risk-group classes (n=4 classes per risk group). Key demographics and population descriptive data are shown in Table 2. The majority of the population was female (n=77, 84%) and White (n=74, 80%). Participants were largely well educated, with all participants having some college degree, and 91% (n=84) having a bachelor's degree or beyond. Additionally, although very few were currently employed (n=15, 16%), the majority (n=55, 60%) of participants were at a high socioeconomic status (>US $100,000 income), with 19 (21%) individuals earning more than US $200,000 per year.

As expected, based upon the group stratification, there was a difference in STEADI fall risk score between groups. Additionally, those at moderate risk (STEADI >3 but <8) were slightly heavier (low risk: mean 64.2, SD 11.3 kg and medium risk: mean 71.6, SD 17.7 kg; *P*=.008) and with a higher BMI (low risk: mean 23.8, SD 3.9 kg/m^2 and medium risk: mean 26.5, SD 5.2 kg/m^2; *P*=.003).

Performance metrics at baseline and following the 12-week intervention are shown in Table 3. There was a difference at baseline between the low- and moderate-risk groups on the 30CS (*P*<.001), TUG (*P*=.002), walking distance at normal speed over 2.5 minutes (*P*<.001), SPPB (*P*<.001), and the OWD measures of posture (normal stance: *P*=.007 and tall stance: *P*=.005). In all of these cases, the moderate-risk group had worse performance compared to the low-risk group. However, strength and balance and some measures of posture (flexiruler and Debrunner kyphometer) were not different by risk category.

Table 2. Demographic data of participants by fall-risk group.

	Total (N=92)	Low risk (n=55)	Moderate risk (n=37)	*P* value
Female, n (%)	77 (84)	47 (85)	30 (81)	.58
Race, n (%)				N/A[a]
Asian	8 (9)	5 (9)	3 (8)	
Black	2 (2)	1 (2)	1 (3)	
Hispanic	3 (3)	1 (2)	2 (5)	
White	74 (80)	45 (82)	29 (78)	
Other	5 (5)	3 (5)	2 (5)	
Age (years), mean (SD)	72.0 (6.6)	71.3 (6.9)	73.0 (6.2)	.25
Height (cm), mean (SD)	163.9 (7.7)	164.0 (6.0)	163.7 (9.9)	.43
Weight (kg), mean (SD)	67.1 (14.6)	64.2 (11.3)	71.6 (17.7)	*.008* [b]
BMI (kg/cm^2), mean (SD)	24.9 (4.6)	23.8 (3.9)	26.5 (5.1)	*.003*
Income (US $ per year), n (%)				
<50,000	12 (13)	4 (7)	8 (22)	.11
>50,000 but <100,000	24 (26)	15 (27)	9 (24)	N/A
>100,000	55 (60)	35 (64)	20 (54)	N/A
STEADI[c] risk score (0-12 range), mean (SD)	3.3 (2.0)	2.0 (1.4)	5.2 (1.3)	*<.001*
Classes attended, mean (SD)	9.7 (3.1)	9.4 (3.5)	10.3 (2.3)	.08

[a]N/A: not applicable.

[b]Values in italics format indicate statistical significance at or below *P*<.05.

[c]STEADI: Stopping Elderly Accidents and Deaths Initiative.

Table 3. Intervention effects on performance-based measures by fall-risk category.

	Total (n=82 except where noted)			Low (n=48 except where noted)			Moderate (n=34 except where noted)		
	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value	Preintervention, mean (SD)	Postintervention, mean (SD)	*P* value
Kyphotic index (%)	11.4 (3.4)	10.7 (3.5)	*.007* [a]	11.4 (3.5)	10.8 (3.3)	.20	11.4 (3.3)	10.5 (3.8)	*.008*
Kyphotic angle (degrees)	40.3 (11.9)	37.7 (13.7)	*.005*	40.3 (10.9)	38 (11.2)	*.02*	40.2 (13.6)	37.3 (17.3)	*.008*
OWD[b]—normal stance (cm)	4.8 (3.8)	3.5 (3.8)	*<.001*	3.9 (3.4)	2.7 (3.2)	*.004*	6.1 (4)	4.7 (4.3)	*<.001*
OWD—tall stance (cm)	2.5 (3.0)	1.6 (2.7)	*<.001*	1.8 (2.7)	0.8 (2.1)	*.002*	3.6 (3.3)	2.8 (2.9)	*.01*
30-Second Chair Stand (repetitions)	12.9 (4.9)	14.3 (5.3)	*.006*	14.7 (4.4)	15.6 (5.1)	.44	10.1 (4.3)	12.3 (5.0)	*.002*
Grip combined (kg)	47.8 (13.8)	47.6 (14.4)	.54	47.9 (14.1)	47.1 (14.1)	.18	47.8 (13.7)	48.1 (15.0)	.60
Isometric leg strength extension (kg) (n=26 with/19 low and 7 moderate)	199 (80.6)	206 (79.1)	*.02*	196.1 (84.6)	193.7 (65.7)	.07	213.4 (62.1)	253.3 (112.6)	.11
Isometric leg strength flexion (kg) (n=26 with/19 low and 7 moderate)	93.9 (42.9)	98.7 (41.1)	.13	94.2 (45)	97.0 (37.7)	.29	92.3 (34.5)	105.2 (56.0)	.20
Isometric trunk strength extension (kg) (n=26 with/19 low and 7 moderate)	70.3 (28.1)	83.0 (40.0)	*.03*	72.5 (28.5)	90.9 (38.0)	*.01*	64.6 (28.1)	61.8 (40.6)	.75
Isometric trunk strength flexion (kg) (n=26 with/19 low and 7 moderate)	43.9 (21.1)	51.8 (20.4)	*<.001*	45.5 (19.8)	54.1 (17.9)	*.002*	39.7 (25.4)	45.9 (26.4)	.16
SPPB[c] (scored from 0 to 12)	11 (1.4)	11.2 (1.2)	.13	11.5 (0.9)	11.5 (1.0)	.40	10.2 (1.6)	10.8 (1.4)	*.02*
Timed Up and Go (seconds)	6.9 (1.7)	6.7 (1.5)	*.004*	6.3 (1.2)	6.1 (1.2)	*.02*	7.9 (1.9)	7.5 (1.6)	.09
Postural sway (cm)	91.4 (31.1)	92.6 (35)	.43	88.4 (27)	86.8 (25.7)	.98	96.7 (37.2)	101.5 (44.8)	.16
Distance during 2.5-minute walk (m)	196.3 (30.7)	193.1 (31.6)	.18	209.2 (29.2)	202.5 (29.9)	.05	179.9 (24.5)	181.0 (33.4)	.75

[a]Values in italics format indicate statistically significant values with a *P* value of less than .05.

[b]OWD: occiput to wall distance.

[c]SPPB: Short Physical Performance Battery.

Balance and Physical Function

On average, at baseline, participants were able to complete slightly more than the minimum number of chair stands (n<12) to be considered free from disability [34]. However, 38 (41%) individuals were below that level. In contrast, on average, participants were well above the minimum threshold to identify disability for the TUG and SPPB (>13.5 seconds and <9, combined score respectively) [35,36], with 0 for TUG and 14 (16%) for SPPB below the threshold. In terms of the effects of the intervention, there were no significant changes in postural sway during static balance stances associated with the SPPB.

However, overall participants significantly improved on the 30CS and TUG (mean change of 1.4 repetitions: *P*=.006 and 0.2 seconds: *P*=.004, respectively). While not always statistically significant, the moderate-risk group improved to a greater degree on both of these assessments than the low-risk group (although both groups improved). Additionally, SPPB performance improved significantly in the moderate-risk category (*P*=.02).

Posture

Although there were no inclusion or exclusion criteria related to posture, there were a meaningful number that displayed suboptimal posture. Indeed, depending upon the measurement tool and associated categorical cut-points used, between 7% and 26% of the total sample were hyperkyphotic (flexiruler: n=7, 8%; kyphometer: n=12, 13%; OWD [normal stance]: n=24, 26%). The majority of postural measures showed statistically significant improvement following the intervention. Indeed, measures of "relaxed" posture showed improvements between 6% (n=5; Debrunner kyphometer) and 27% (n=22; OWD). When an "as tall as possible" posture was adopted for the OWD, the improvements were greater at ~36% change from baseline. These changes are even more impressive when individuals who had no opportunity for improvement are removed (ie, individuals who scored 0 on OWD at baseline). Specifically, when the 18 individuals standing normally and 47 as tall as possible who scored 0 are removed, overall change scores with normal stance were 1.7 cm, and with tall stance were 2.2 cm.

Muscular Strength

We did not observe significant differences in grip strength (*P*=.54) However, as mentioned earlier, 30CS is a marker both of functional movement but also muscular strength (and endurance) and had significant changes across the population (*P*=.006) Finally, in the subgroup that completed isometric testing, there was a significant improvement in knee extension (*P*=.02) and both trunk extension and flexion (*P*=.03 and <.001, respectively).

Aerobic Fitness

There were no significant changes in measures of aerobic fitness measured by the 2.5-minute walk at normal speed.

Associations Between Class Attendance and Functional Measures

When exploring associations between change scores and rates of attendance, the TUG (*r*=–0.255; *P*=.03) and 30CS (*r*=0.27; *P*=.02) were significantly correlated; however, there were no significant associations between class attendance and changes in physical function on the majority of other variables (*P*>.05). Regression analysis controlling for sex, age, baseline measurement, and number of classes attended indicated that attendance was not a significant predictor of change in TUG, although including it did slightly improve the model's predictive value (from *R*=0.45 to 0.49). However, attendance was a significant predictor of improvement in the 30CS (unstandardized β=.502; *P*=.02). This indicates that for each class attended, there was a 0.5 increase in the change score in terms of number of repetitions. Model summary and individual contributions of variables and covariates are shown in Tables 4 and 5.

Table 4. Linear regression analysis of functional fitness and attendance change in the Timed Up and Go[a].

Predictors	Unstandardized β (SE)	Standardized β coefficient	*t* test (*df*=72)	*P* value
Constant	–.382 (0.981)	N/A[b]	–0.389	.70
Sex[c]	–.066 (0.19)	–0.036	–0.346	.73
Chronological age (years)	.031 (0.012)	0.297	2.657	*.01*[d]
Baseline TUG[e] (seconds)	–.174 (0.047)	–0.417	–3.72	*<.001*
Attendance	–.073 (0.04)	–0.191	–1.81	.07

[a]Model summary: *R*=0.485; R^2=0.235; adjusted R^2=0.193; standard error of the estimate=0.6118; *P*=.07.

[b]N/A: not applicable.

[c]Male: n=1 and female: n=2.

[d]Values in italics format indicate significance at, or below, the level of .05.

[e]TUG: Timed Up and Go.

Table 5. Linear regression analysis of functional fitness and attendance change in the 30-second Chair Stand[a].

Predictors	Unstandardized β (SE)	Standardized β coefficient	*t* test (*df*=72)	*P* value
Constant	6.172 (5.689)	N/A[b]	1.085	.28
Sex[c]	–.785 (0.991)	–0.089	–0.792	.43
Chronological age (years)	–.096 (0.06)	–0.185	–1.6	.11
Baseline 30-second Chair Stand	–.156 (0.079)	–0.227	–1.978	.05
Attendance	.502 (0.21)	0.266	2.393	*.02*[d]

[a]Model summary: R=0.374; R^2=0.14; adjusted R^2=0.092; standard error of the estimate=3.173; P=.02.

[b]N/A: not applicable.

[c]Male: n=1 and female: n=2.

[d]Values in italics format indicate significance at, or below, the level of .05.

Adverse Events

Based upon weekly tracking, there were several minor adverse events throughout the intervention period. However, none were related to the exercise intervention, and only 1 minor event was possibly related to the measurements. Additionally, there were no falls associated with the intervention, either while supervised during class or performed on nonclass days.

Discussion

Principal Findings

In this 12-week, iterative, web-based program, we found that older adults experienced an improvement in several measures widely used to evaluate fall risk and mobility-related disability [7,28,34]. While the magnitude of most of the changes was small, they were primarily consistent in the expected direction, which is not surprising because the program was designed to be largely instructional rather than a progressive training program. Instructors emphasized proper form and execution while maintaining good posture. Increases in the intensity of resistance exercises and challenges to balance and mobility exercises were only introduced once instructors observed consistent execution of movements. Importantly, no injuries were sustained during the remote delivery of our program or during the recommended home exercise practice.

Most muscular strength and physical function measures showed favorable changes in both the low- and moderate-risk groups, although not all results reached statistical significance. Nevertheless, the observed changes have clinical relevance. For instance, the intervention induced meaningful improvements in measures commonly used in clinical settings, such as the 30CS and TUG tests. Notably, many improvements varied by risk level, with greater gains seen in those at higher fall risk, who arguably had more to gain from the intervention. Specifically, participants in the moderate-risk group completed an additional 2 chair rises (a 21% improvement) by the end of the 12 weeks. Given that both leg strength and muscular endurance are essential for scoring well on the 30CS, the observed improvement in chair stands may be attributed to the focused practice of squat and related exercises in class as well as encouragement for similar home practice. While upper body or arm strength is important to overall good health and function, they do not substantially contribute to fall prevention and were therefore not heavily emphasized in our home practice recommendations. Similarly, the lack of significant changes in aerobic capacity was expected based on the program's goals and the unique challenges of aerobic activities conducted via a Zoom format (Zoom Video Communications).

Additionally, these findings also contextualize participant gains by providing deeper phenotypical measurements related to posture, more sophisticated measures of strength, and balance. For example, there were positive changes recorded in measures of thoracic kyphosis and OWD across both risk groups. This is important because posture is often overlooked in many group exercise programs for both younger and older adults. In this study population, posture measured as the kyphotic angle using a Debrunner kyphometer decreased 2-3 degrees (P=.02 and .07 for low- and moderate-risk groups, respectively), while the OWD decreased between 1 and 1.4 cm (all P<.05) among participants in both groups over the once-weekly, 12-week exercise program. Considering that thoracic kyphosis typically worsens steadily over time beginning as early as age 40 years [37-39], observing a reversal of this condition within just 12 weeks of targeted exercise is highly encouraging. This is particularly meaningful, given the observed links between hyperkyphosis and both fall risk [26,40] and physical function [41,42].

Although there were no significant differences observed in grip strength, it is likely that the prescribed exercises contributed to an enhancement of the participant's back strength [43] and chest flexibility, which helped alleviate the tightness in chronically shortened pectoral muscles [26]. These improvements not only reflected positively in participants' attempts at maintaining "good" posture (as measured by OWD while standing tall) but also in their "normal" posture as assessed by various tools. These changes in both "good" and "normal" posture likely stemmed from the early introduction of postural control as a foundational element of the program as well as ongoing reminders throughout the course for participants to practice maintaining good posture during daily activities. It is worth noting that postural control is seldom included in group exercise training programs, making this approach particularly novel and engaging, thereby providing participants with more opportunities to learn and improve compared to more traditional exercise modalities such as aerobic activities or strength training.

While there have been some similar findings regarding the possibility of providing digital exercise instruction to older adults [17-19], this is, to our knowledge, the first study to robustly measure a large population of older adults across multiple metrics of physical health and function. Our findings clearly demonstrate the potential to deliver an effective fall prevention program through a technological interface. The program was well-received, as indicated by an attendance rate exceeding 80% (mean attendance 9.7 of 12), and no adverse events were reported in association with the exercise regimen. Additionally, the observed improvements in leg strength, postural control, and overall mobility, both independently and synergistically, show promise for reducing falls among older adults, including those at substantial risk.

Limitations

Despite several strengths of this study, certain limitations must be acknowledged. Chief among them is the absence of a control group, which limits our ability to ascertain whether observed changes were influenced by external factors. A control group would also help clarify the significance of nonsignificant changes observed in some metrics, particularly in light of the expected performance declines over time. Additionally, although there were improvements in most metrics of strength and posture, balance as measured in this context was not improved. This may be a function of the method of measuring static balance (ie, postural sway during the SPPB). However, it may be that the program's focus on balance was not sufficient to induce changes over 12 weeks. Furthermore, unlike many exercise interventions for adults of all ages, we did not emphasize cardiovascular health other than recommending daily general PA. While recognizing the importance of aerobic activity for overall health and longevity, instructional focus on this area was minimal; encouragement for at-home exercise emphasized practicing postural, strength, and balance-based exercises. Finally, the relatively affluent nature of the population limits the generalizability of both the results and the ability to disseminate this program to a wider population.

With the success of this program and the limitations noted earlier in mind, future research should explore the possibility of deploying this intervention in populations that are (1) of a lower overall socioeconomic status and (2) more remote from the location of intervention deployment. In addition, developing these materials into other languages in a culturally appropriate manner offers substantial potential to expand the reach to other populations who would benefit from the opportunity to receive at-home fall-risk training.

Conclusions

The 12-week SF program improved physical function and posture in ways consistent with reduced fall risk. The real-time instruction also helps ensure program safety and adherence. As such, this program shows promise in digitally delivering needed public health interventions targeting fall risk among older adults.

Acknowledgments

Study data were collected and managed using REDCap (Research Electronic Data Capture) electronic data capture tools hosted at the University of California, San Diego (UCSD). The REDCap software system provided by the UCSD Clinical and Translational Research Center is supported by (award UL1TR001442) the National Center for Research Resources. The authors would like to thank Ms Ashley Gauthie and Ms Sharmaine Olivarez for their work in managing the technical elements of Zoom-based intervention deployment. The authors would also like to thank their funder, the UCSD Clinical and Translation Research Center.

Data Availability

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

Conflicts of Interest

None declared.

References

1. The graying of America: more older adults than kids by 2035. United States Census Bureau. 2018. URL: https://www.census.gov/library/stories/2018/03/graying-america.html [accessed 2024-10-03]
2. Stevens J. A CDC compendium of effective fall interventions: what works for community-dwelling older adults. In: Injury Prevention, 2nd Edition. Scotts Valley, CA: CreateSpace Independent Publishing; 2010.
3. Lundin-Olsson L, Nyberg L, Gustafson Y. Attention, frailty, and falls: the effect of a manual task on basic mobility. J Am Geriatr Soc 1998;46(6):758-761. [doi: 10.1111/j.1532-5415.1998.tb03813.x] [Medline: 9625194]
4. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med 1988;319(26):1701-1707. [doi: 10.1056/NEJM198812293192604] [Medline: 3205267]
5. Baloh RW, Corona S, Jacobson KM, Enrietto JA, Bell T. A prospective study of posturography in normal older people. J Am Geriatr Soc 1998;46(4):438-443. [doi: 10.1111/j.1532-5415.1998.tb02463.x] [Medline: 9560065]
6. Espinosa de Los Monteros K, Gallo LC, Elder JP, Talavera GA. Individual and area-based indicators of acculturation and the metabolic syndrome among low-income Mexican American women living in a border region. Am J Public Health 2008;98(11):1979-1986 [FREE Full text] [doi: 10.2105/AJPH.2008.141903] [Medline: 18799765]

7. Gill TM. Assessment of function and disability in longitudinal studies. J Am Geriatr Soc 2010;58(Suppl 2):S308-S312 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.02914.x] [Medline: 21029059]
8. Stevens JA, Phelan EA. Development of STEADI: a fall prevention resource for health care providers. Health Promot Pract 2013;14(5):706-714 [FREE Full text] [doi: 10.1177/1524839912463576] [Medline: 23159993]
9. Sarmiento K, Lee R. STEADI: CDC's approach to make older adult fall prevention part of every primary care practice. J Safety Res 2017;63:105-109 [FREE Full text] [doi: 10.1016/j.jsr.2017.08.003] [Medline: 29203005]
10. US Preventive Services Task Force, Grossman DC, Curry SJ, Owens DK, Barry MJ, Caughey AB, et al. Interventions to prevent falls in community-dwelling older adults: US Preventive Services Task Force recommendation statement. JAMA 2018;319(16):1696-1704 [FREE Full text] [doi: 10.1001/jama.2018.3097] [Medline: 29710141]
11. Belza B, Shumway-Cook A, Phelan EA, Williams B, Snyder SJ, LoGerfo JP. The effects of a community-based exercise program on function and health in older adults: the EnhanceFitness program. J Appl Gerontol 2006 Aug 01;25(4):291-306. [doi: 10.1177/0733464806290934]
12. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014;12(1-2):10-20 [FREE Full text] [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
13. Pelletier CA, White N, Duchesne A, Sluggett L. Barriers to physical activity for adults in rural and urban Canada: a cross-sectional comparison. SSM Popul Health 2021;16:100964 [FREE Full text] [doi: 10.1016/j.ssmph.2021.100964] [Medline: 34841038]
14. Kakara R, Bergen G, Burns E, Stevens M. Nonfatal and fatal falls among adults aged ≥65 years—United States, 2020–2021. MMWR Morb Mortal Wkly Rep 2023;72(35):938-943 [FREE Full text] [doi: 10.15585/mmwr.mm7235a1] [Medline: 37651272]
15. Mace RA, Mattos MK, Vranceanu A. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022;12(12):1102-1105 [FREE Full text] [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
16. Jacobson CL, Foster LC, Arul H, Rees A, Stafford RS. A digital health fall prevention program for older adults: feasibility study. JMIR Form Res 2021;5(12):e30558 [FREE Full text] [doi: 10.2196/30558] [Medline: 34837492]
17. Yerlikaya T, Öniz A, Özgüren M. The effect of an interactive tele rehabilitation program on balance in older individuals. Neurol Sci Neurophys 2021;38(3):180-186 [FREE Full text] [doi: 10.4103/nsn.nsn_91_21]
18. Ing JBM, Tan MP, Whitney J, Tiong IK, Singh DKA. Acceptability, feasibility, and effectiveness of WE-SURF™: a virtual supervised group-based fall prevention exercise program among older adults. Aging Clin Exp Res 2024;36(1):125 [FREE Full text] [doi: 10.1007/s40520-024-02759-x] [Medline: 38836944]
19. Langeard A, Bigot L, Maffiuletti NA, Moussay S, Sesboüé B, Quarck G, et al. Non-inferiority of a home-based videoconference physical training program in comparison with the same program administered face-to-face in healthy older adults: the MOTION randomised controlled trial. Age Ageing 2022;51(3):afac059. [doi: 10.1093/ageing/afac059] [Medline: 35290431]
20. Blioumpa C, Karanasiou E, Antoniou V, Batalik L, Kalatzis K, Lanaras L, et al. Efficacy of supervised home-based, real time, videoconferencing telerehabilitation in patients with type 2 diabetes: a single-blind randomized controlled trial. Eur J Phys Rehabil Med 2023;59(5):628-639 [FREE Full text] [doi: 10.23736/S1973-9087.23.07855-3] [Medline: 37350165]
21. Terkes N, Aksu NT, Yamac SU. The effect of an online-supervised exercise program in older people with diabetes on fasting blood sugar, psychological resilience and quality of life: a double blind randomised controlled trial. Int J Older People Nurs 2023;18(5):e12564. [doi: 10.1111/opn.12564] [Medline: 37559428]
22. Li F, Harmer P, Fitzgerald K, Winters-Stone K. A cognitively enhanced online Tai Ji Quan training intervention for community-dwelling older adults with mild cognitive impairment: a feasibility trial. BMC Geriatr 2022;22(1):76 [FREE Full text] [doi: 10.1186/s12877-021-02747-0] [Medline: 35078407]
23. Pitre V, Sénéchal M, Bouchard DR. Comparing online and in-person delivery of a fall prevention exercise program for older adults. J Aging Phys Act 2024;32(1):55-61. [doi: 10.1123/japa.2023-0070] [Medline: 37741634]
24. Granet J, Peyrusqué E, Ruiz F, Buckinx F, Abdelkader LB, Dang-Vu TT, et al. Web-based physical activity interventions are feasible and beneficial solutions to prevent physical and mental health declines in community-dwelling older adults during isolation periods. J Gerontol A Biol Sci Med Sci 2023;78(3):535-544 [FREE Full text] [doi: 10.1093/gerona/glac127] [Medline: 35675174]
25. Kado DM, Huang M, Nguyen CB, Barrett-Connor E, Greendale GA. Hyperkyphotic posture and risk of injurious falls in older persons: the Rancho Bernardo Study. J Gerontol A Biol Sci Med Sci 2007;62(6):652-657. [doi: 10.1093/gerona/62.6.652] [Medline: 17595423]
26. van der Jagt-Willems HC, de Groot MH, van Campen JPCM, Lamoth CJC, Lems WF. Associations between vertebral fractures, increased thoracic kyphosis, a flexed posture and falls in older adults: a prospective cohort study. BMC Geriatr 2015;15:34 [FREE Full text] [doi: 10.1186/s12877-015-0018-z] [Medline: 25888399]
27. Rubenstein LZ, Vivrette R, Harker JO, Stevens JA, Kramer BJ. Validating an evidence-based, self-rated Fall Risk Questionnaire (FRQ) for older adults. J Safety Res 2011;42(6):493-499. [doi: 10.1016/j.jsr.2011.08.006] [Medline: 22152267]

28. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
29. Podsiadlo D, Richardson S. The timed "Up and Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
30. Wiyanad A, Chokphukiao P, Suwannarat P, Thaweewannakij T, Wattanapan P, Gaogasigam C, et al. Is the occiput-wall distance valid and reliable to determine the presence of thoracic hyperkyphosis? Musculoskelet Sci Pract 2018;38:63-68. [doi: 10.1016/j.msksp.2018.09.010] [Medline: 30278368]
31. Milne JS, Lauder IJ. The relationship of kyphosis to the shape of vertebral bodies. Ann Hum Biol 1976;3(2):173-179 [FREE Full text] [doi: 10.1080/03014467600001281] [Medline: 1275439]
32. Ohlén G, Spangfort E, Tingvall C. Measurement of spinal sagittal configuration and mobility with Debrunner's kyphometer. Spine (Phila Pa 1976) 1989;14(6):580-583. [doi: 10.1097/00007632-198906000-00006] [Medline: 2749372]
33. Rossman I. Normal human aging: the Baltimore Longitudinal Study of Aging. JAMA 1986;255(7):960. [doi: 10.1001/jama.1986.03370070114046]
34. Rikli R, Jones C. Functional fitness normative scores for community-residing older adults, ages 60-94. J Aging Phys Act 1999;7(2):129-161 [FREE Full text] [doi: 10.1123/japa.7.2.129]
35. Rose D, Jones C, Lucchese N. Predicting the probability of falls in community-residing older adults using the 8-foot up-and-go: a new measure of functional mobility. J Aging Phys Act 2002;10(4):466-475 [FREE Full text] [doi: 10.1123/japa.10.4.466]
36. de Fátima Ribeiro Silva C, Ohara DG, Matos AP, Pinto ACPN, Pegorari MS. Short Physical Performance Battery as a measure of physical performance and mortality predictor in older adults: a comprehensive literature review. Int J Environ Res Public Health 2021;18(20):10612 [FREE Full text] [doi: 10.3390/ijerph182010612] [Medline: 34682359]
37. Fon GT, Pitt MJ, Thies AC. Thoracic kyphosis: range in normal subjects. Am J Roentgenol 1980;134(5):979-983. [doi: 10.2214/ajr.134.5.979] [Medline: 6768276]
38. Katzman WB, Wanek L, Shepherd JA, Sellmeyer DE. Age-related hyperkyphosis: its causes, consequences, and management. J Orthop Sports Phys Ther 2010;40(6):352-360 [FREE Full text] [doi: 10.2519/jospt.2010.3099] [Medline: 20511692]
39. Ensrud KE, Black DM, Harris F, Ettinger B, Cummings SR. Correlates of kyphosis in older women. The Fracture Intervention Trial Research Group. J Am Geriatr Soc 1997;45(6):682-687. [doi: 10.1111/j.1532-5415.1997.tb01470.x] [Medline: 9180660]
40. McDaniels-Davidson C, Davis A, Wing D, Macera C, Lindsay SP, Schousboe JT, et al. Kyphosis and incident falls among community-dwelling older adults. Osteoporos Int 2018;29(1):163-169 [FREE Full text] [doi: 10.1007/s00198-017-4253-3] [Medline: 29018904]
41. Katzman WB, Vittinghoff E, Ensrud K, Black DM, Kado DM. Increasing kyphosis predicts worsening mobility in older community-dwelling women: a prospective cohort study. J Am Geriatr Soc 2011;59(1):96-100 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03214.x] [Medline: 21198460]
42. Katzman WB, Huang M, Lane NE, Ensrud KE, Kado DM. Kyphosis and decline in physical function over 15 years in older community-dwelling women: the study of osteoporotic fractures. J Gerontol A Biol Sci Med Sci 2013;68(8):976-983 [FREE Full text] [doi: 10.1093/gerona/glt009] [Medline: 23633167]
43. Katzman WB, Vittinghoff E, Lin F, Schafer A, Long RK, Wong S, et al. Targeted spine strengthening exercise and posture training program to reduce hyperkyphosis in older adults: results from the study of hyperkyphosis, exercise, and function (SHEAF) randomized controlled trial. Osteoporos Int 2017;28(10):2831-2841 [FREE Full text] [doi: 10.1007/s00198-017-4109-x] [Medline: 28689306]

Abbreviations

30CS: 30-second Chair Stand
CDC: Centers for Disease Control and Prevention
OWD: occiput to wall distance
PA: physical activity
SF: Strong Foundations
SPPB: Short Physical Performance Battery
STEADI: Stopping Elderly Accidents and Deaths Initiative
TUG: Timed Up and Go

Edited by M Gray; submitted 21.11.24; peer-reviewed by D Kunkel, E Okoyeocha; accepted 02.01.25; published 26.02.25.

Please cite as:
Wing D, Nichols JF, Barkai HS, Culbert O, Moreno D, Higgins M, O'Brien A, Perez M, Davey H, Moran R
Building Strong Foundations: Nonrandomized Interventional Study of a Novel, Digitally Delivered Fall Prevention Program for Older Adults
JMIR Aging 2025;8:e68957
URL: https://aging.jmir.org/2025/1/e68957
doi:10.2196/68957
PMID:

Original Paper

Development of a Predictive Dashboard With Prescriptive Decision Support for Falls Prevention in Residential Aged Care: User-Centered Design Approach

S Sandun Malpriya Silva[1], BSc, PhD; Nasir Wabe[1], BPharm, MScClinPharm, MScEpid, PhD; Amy D Nguyen[1], BMedSc, PhD; Karla Seaman[1], BPharm, MClinEpid, PhD; Guogui Huang[1], BA, MA, PhD; Laura Dodds[2], BA, MPH; Isabelle Meulenbroeks[1], BPhysio, MPH, PhD; Crisostomo Ibarra Mercado[1], BSN, MPH; Johanna I Westbrook[1], BAppSc, GradDipAppEpid, MHA, PhD

[1]Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Macquarie University, North Ryde, NSW, Australia

[2]The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, New South Wales, Australia

Corresponding Author:
S Sandun Malpriya Silva, BSc, PhD
Centre for Health Systems and Safety Research
Australian Institute of Health Innovation
Macquarie University
75 Talavera Road
North Ryde, NSW, 2113
Australia
Phone: 61 29850 4045
Email: sandunmalpriya.silva@mq.edu.au

Abstract

Background: Falls are a prevalent and serious health condition among older people in residential aged care facilities, causing significant health and economic burdens. However, the likelihood of future falls can be predicted, and thus, falls can be prevented if appropriate prevention programs are implemented. Current fall prevention programs in residential aged care facilities rely on risk screening tools with suboptimal predictive performance, leading to significant concerns regarding resident safety.

Objective: This study aimed to develop a predictive, dynamic dashboard to identify residents at risk of falls with associated decision support. This paper provides an overview of the technical process, including the challenges faced and the strategies used to overcome them during the development of the dashboard.

Methods: A predictive dashboard was co-designed with a major residential aged care partner in New South Wales, Australia. Data from resident profiles, daily medications, fall incidents, and fall risk assessments were used. A dynamic fall risk prediction model and personalized rule-based fall prevention recommendations were embedded in the dashboard. The data ingestion process into the dashboard was designed to mitigate the impact of underlying data system changes. This approach aims to ensure resilience against alterations in the data systems.

Results: The dashboard was developed using Microsoft Power BI and advanced R programming by linking data silos. It includes dashboard views for those managing facilities and for those caring for residents. Data drill-through functionality was used to navigate through different dashboard views. Resident-level change in daily risk of falling and risk factors and timely evidence-based recommendations were output to prevent falls and enhance prescriptive decision support.

Conclusions: This study emphasizes the significance of a sustainable dashboard architecture and how to overcome the challenges faced when developing a dashboard amid underlying data system changes. The development process used an iterative dashboard co-design process, ensuring the successful implementation of knowledge into practice. Future research will focus on the implementation and evaluation of the dashboard's impact on health processes and economic outcomes.

International Registered Report Identifier (IRRID): RR2-https://doi.org/10.1136/bmjopen-2021-048657

(JMIR Aging 2025;8:e63609) doi:10.2196/63609

KEYWORDS

falls prevention; dashboard architecture; predictive; sustainability; challenges; decision support; falls; aged care; geriatric; older adults; economic burden; prevention; electronic health record; EHR; intervention; decision-making; patient safety; risks; older people; monitoring

Introduction

Falls are the second leading cause of unintentional injury deaths worldwide, with those older than 60 years having the highest risk of death or serious injury [1]. People living in residential aged care facilities (RACFs; ie, nursing homes and care homes) are three times more likely to fall than their community-dwelling peers and are ten times more likely to sustain fall-related injuries [2]. Studies have found that falls are often predictable and preventable and that the health and economic burden of falls could be minimized if effective prevention strategies are implemented [3-5].

Falls prevention is an issue of major importance in RACFs. Various fall risk assessment tools (FRATs) are used by RACF staff internally to identify residents at greatest risk of falls [6]. However, when assessed, the predictive performance of these tools has generally shown to be poor [6,7]. Moreover, FRATs are most often completed when residents are admitted to a facility or after a fall incident, and thus, only represent data collected at a single point in time, resulting in a static prediction that does not reflect changing falls risk factors (eg, if a resident's medication use changes) [6].

As a growing number of RACFs implement electronic health record (EHR) systems, new opportunities have emerged to develop a personalized, dynamic approach to predicting residents' fall risk by taking advantage of multiple potential contributory factors [8]. Some studies have integrated routinely collected EHR data including vital signs into the development of fall prediction tools through the application of machine learning models [9]. Despite these tools exhibiting superior performance in contrast to conventional FRATs, the risk predictions generated are often difficult to interpret and apply in practice. Thus, relying solely on these models will not enhance decision-making to provide a personalized care strategy for fall prevention and management. Health dashboards that can integrate these predictive models in combination with visualization features may be a more effective intervention for supporting staff in RACFs.

A dashboard can be defined as an at-a-glance real-time or near real-time processing interface, showing a graphical presentation of the current or recent status, along with historical trends or organizational key performance indicators to support decision-making [10]. There have been some attempts to develop a dashboard for fall prevention purposes globally. The falls dashboard, developed by the Network of Patient Safety Databases and managed by the US Department of Health and Human Services, collects nonidentifiable patient-related fall incident data from health care service providers. These data are then aggregated to a national level, analyzed, and visualized using the Network of Patient Safety Database dashboards and chartbooks to identify and track patient safety risks nationally [11]. Similarly, the New Zealand Health Quality and Safety Commission has established a falls and fracture outcomes dashboard to help the health sector evaluate the benefits of the services provided to older people [12]. Such aggregated dashboards are helpful for monitoring, although they offer limited assistance in predicting and preventing falls in routine care settings. To the best of our knowledge, no previous study has reported the development of a dashboard designed to predict fall risks for individual residents and aid decision-making in preventing and managing falls. As part of a comprehensive research program aimed at enhancing fall prevention and the well-being of older adults in RACFs [8], this paper details the design and development process of a predictive analytics dashboard with prescriptive decision support for fall risk prediction, prevention, and management. We discuss the challenges encountered and the strategies used to overcome them during its development.

Methods

Participants

This project was conducted in collaboration with a large, aged care provider that operates 24 RACFs in Sydney, Australia. The development process of the dashboard spanned from January 1, 2022, to December 31, 2023. For the dashboard views and study sample, this study used data from both respite and permanent residents who were present in a RACF between January 1, 2021, and December 31, 2021. The dashboard was designed using participant data that was extracted onto Excel sheets from the provider. After the development, the dashboard was connected to the provider's live data sources through a data lake.

Ethical Considerations

All methods and analyses in this study were conducted in accordance with the principles of the Declaration of Helsinki. The study received ethics approval from the Macquarie University Human Research Ethics Committee (52019614412614). As a retrospective cohort study using deidentified, routinely collected aged care data, a waiver of informed consent was granted by the committee, in line with the Australian National Statement on Ethical Conduct in Human Research (Sect. 3.3.14(a)), which allows for the use of previously collected deidentified data.

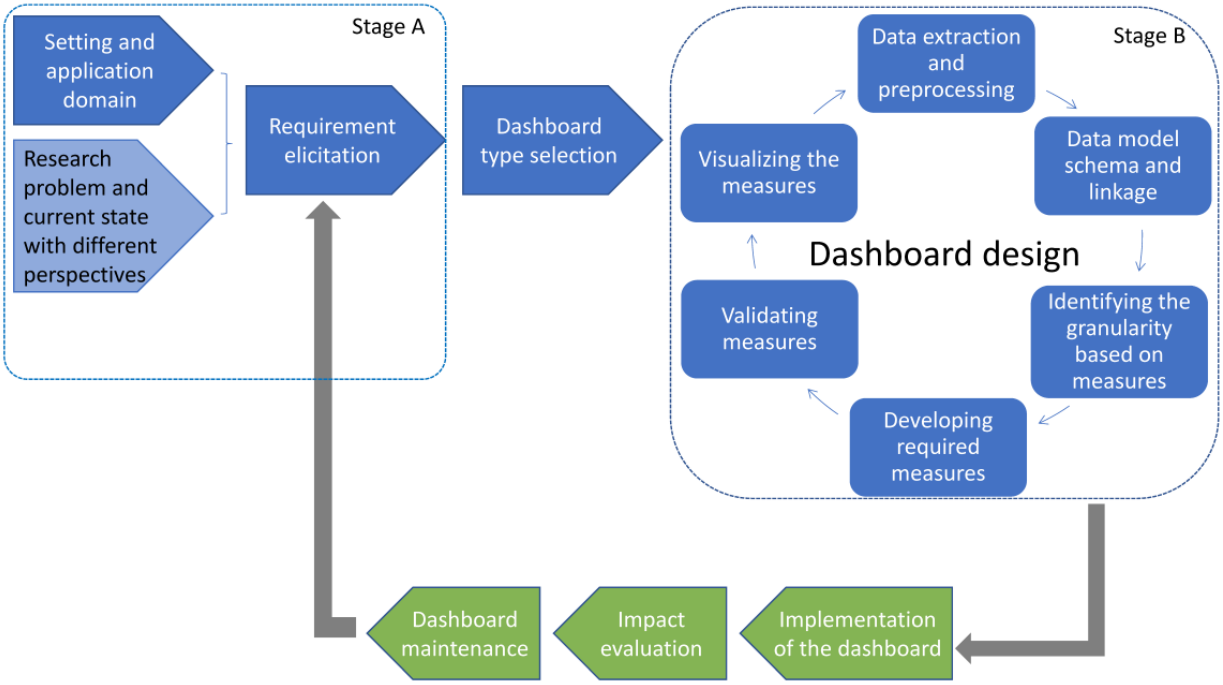
Dashboard Development Process

Overview

We followed an iterative process to co-design the dashboard, with intended dashboard end users (ie, RACF staff, general practitioners, and consumers at RACFs), along with domain experts in the field of falls prevention and management. Figure 1 highlights the dashboard development process followed during the study. The stakeholders were involved continuously in Stage A and Stage B through regular discussions, interviews, and

workshops, which will be described in detail elsewhere. The steps colored in blue are described in this paper and the final steps relating to implementation and evaluation colored in green are in progress and will be reported in subsequent papers.

Figure 1. The dashboard development process.



Setting and Application Domain

Domain knowledge can be defined as the knowledge about the environment in which the proposed system will operate [13]. Domain knowledge was obtained from the aged care provider through discussions and analysis of data related to falls incidents and the Peninsula Health Falls Risk Assessment Tool (PH-FRAT) risk assessments (ie, to determine the current falls rate across facilities, the characteristics of falls incidents and contributing factors listed in falls incidents). PH-FRAT was developed in 1999 by Peninsula Health in Victoria, Australia, and is a validated and user-friendly tool for screening and assessing fall risk, as well as managing strategies to reduce this risk [14]. Information about predictors of falls and strategies used to prevent falls in the RACFs was obtained by systematically reviewing published literature on predictive models for fall prevention, and the effectiveness of fall prevention interventions [15,16]. The identified predictors of falls in older people have been incorporated into the development of the embedded predictive model, which is reported elsewhere [17].

Research Problem and Current State With Different Perspectives

Considering the suboptimal predictive performance of current tools [6], there is a need for a dashboard embedded with a dynamic predictive model and decision support (research problem). Moreover, falls prevention and management requires action at different levels of an organization's hierarchy from management to clinical staff. Therefore, it was important to gather different perspectives. For this purpose, a series of discussions were held between researchers and RACF stakeholders across a 2-stage process (Figure 1). In Stage A discussions, different perspectives on falls prevention and management in RACFs were gathered while identifying the dashboard stakeholders and their information requirements. In Stage B, dashboard stakeholders engaged in rounds of iterative feedback workshops to support the dashboard designing process. A qualitative analysis of these interviews will be reported elsewhere.

Requirement Elicitation

The requirement elicitation stage was based on three fundamental tasks: (1) identifying the individual stakeholders; (2) eliciting stakeholder requirements depending on the role of the stakeholder; and (3) integrating, refining, and organizing the requirements identified [18].

In Stage A, baseline interviews were conducted to understand the current utilization of health care data and information by potential dashboard end users. The aim was to identify the information requirements essential for enhancing decision-making in a clinical dashboard. In this process, the individual stakeholders and information providers of the dashboard were identified along with their high-level user expectations [10]. This also facilitated a precise understanding of the individuals with different perspectives of falls prevention and management, who would be accessing each dashboard view. The information collected from these interviews was used to define measures that should be generated in the dashboard backend.

From these identified measures, the designers then evaluated the variability of information and organized the information requirements based on the level of reporting in the dashboard

(ie, resident level, facility level, or organizational level). This further supported the dashboard design stage in arranging visual plots and tables within views.

Dashboard Type Selection

Power BI (PBI; Microsoft Corp) was selected on the grounds of digital data visualization, easy implementation at the user end, and stable R programming integration for advanced statistical analysis and plots. In this step, the initial prototype of the dashboard was created and shared with our partner organization and refined through collaborative efforts.

Dashboard Design

Overview

The aim of this step was to meet the user requirements for preventing and managing falls by effectively presenting resident data and fall risk information to support informed decision-making. As in Figure 1, this step included an iterative process that included data extraction, preprocessing, and linkage along with identification of the granularity of data, development, validation, and visualization of the measures on the front end.

Data Extraction and Preprocessing

Access to four datasets was obtained: the resident profile (includes residents' demographics and admission information); all medications administered to residents; the organization's PH-FRATs (this aged care providers used the PH-FRAT) to obtain information related to falls risk assessments; and incident dataset (includes information related to all fall incidents and pressure injuries; Table S1 in Multimedia Appendix 1).

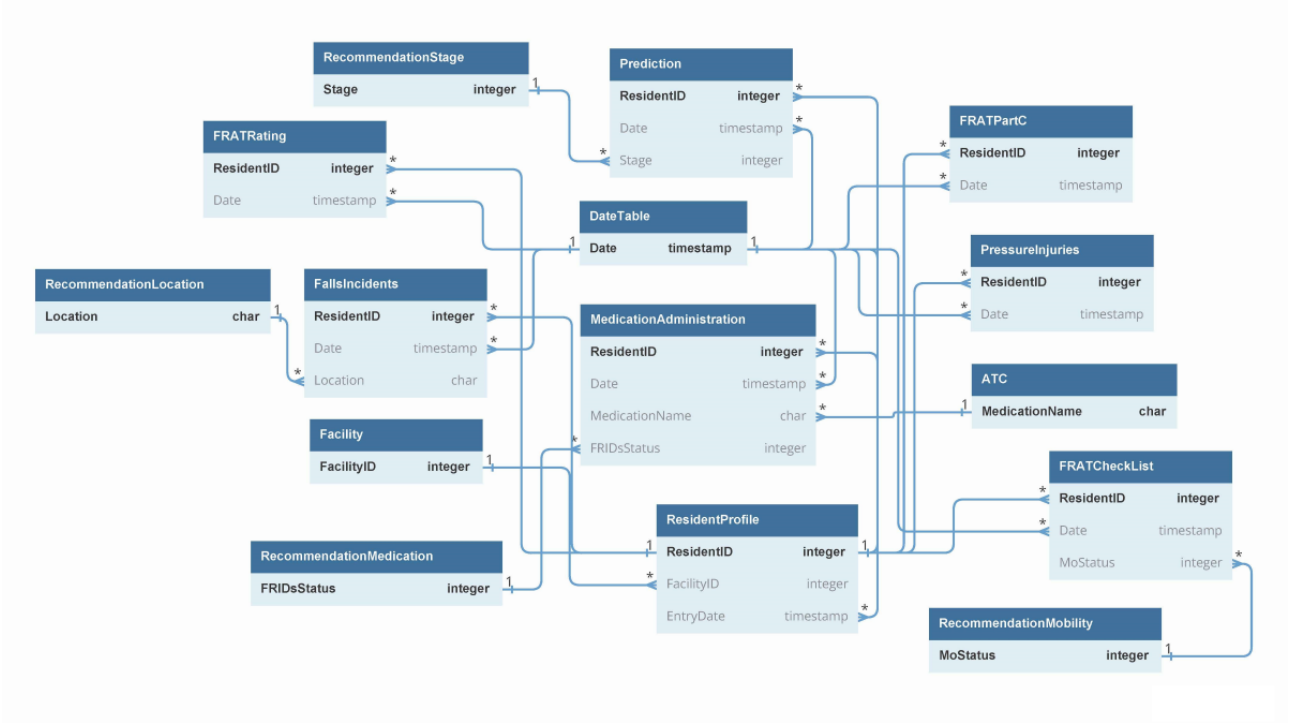
The profile dataset included a free text field that reported the comorbidities along with other special needs of the patient at admission (ie, health status). From this field, comorbidities present at admission were identified using the R-programmed version of the "aged care health status algorithm" [19] within PBI. The algorithm identifies the health conditions using free text fields from EHRs [20]. All medications in the dataset were coded using the Anatomical Therapeutic Chemical codes [21]. These datasets were then linked in the dashboard backend.

The extracted datasets for the study period underwent an external analysis for this study using R programming language (version 4.3.3; R Core Team) and were also subsequently used for its intended purpose within PBI. A descriptive analysis of the admission-related information from the resident profile dataset is reported appropriately (Table S2 in Multimedia Appendix 1). Resident characteristics from the resident profile dataset, FRAT dataset, and daily medication administration data are recorded in Table S3 in Multimedia Appendix 1.

Data Model Schema and Linkage

A star schema was adopted as much as possible where the fact table is centered and surrounded by the multiple dimension tables. The resident profile dataset was used as the main fact table that contains quantitative information (facts) about the residents [22]. Other datasets were used as dimension tables, which describe the events that residents experience during their stay [22]. The data tables were then linked as presented in the simplified form of the database schema as shown in Figure 2. The real schema on the backend of the dashboard is an extension of this and is more advanced.

Figure 2. Simplified view of the data model schema. ATC: Anatomical Therapeutic Chemical code; FRAT: falls risk assessment tool.



Identifying Granularity Based on Measures

To determine the granularity of the data tables, we considered the dashboard end users, information requirements (measures), and level of reporting that were preidentified. This step was done in parallel or a back-and-forth manner with the data model schema (ie, preceding step) and development of required measures (ie, proceeding step). A higher granularity was

preferred for most of the datasets to maintain sustainable analytical power to answer new user requirements within the designing stage [23]. Subsequently, the data tables were transformed to simply calculate the measures, or measure calculations were adjusted in a way to meet the form and granularity in the extracted tables. In certain instances, highly granular data tables were summarized to a lower-level granularity to support the prediction model and avoid intricate calculations during measure development.

Developing the Required Measures

In this step, measures were calculated on the transformed tables using Data Analysis Expression (DAX). Measures related to some figures such as funnel plots were created using embedded R programming in the PBI backend.

Validating the Measures

To ensure the accuracy and reliability of the information presented, the calculated measures were validated with provider-extracted datasets. The user input filters on each view were also tested during the validation process to identify whether there were any errors. If errors were identified in the data, remedial action was taken after a thorough examination of the underlying causes, in data extraction, data linkage, data preprocessing, or DAX formula.

Visualizing the Measures

The calculated measures were appropriately visualized using statistical plots, tables, and responsive figures on the dashboard, which comprises several dashboard views intended for various users at different levels in the hierarchy. The above-mentioned steps within the dashboard design stage were enclosed in a co-design process as explained below.

Dashboard Co-Design and Work Process Features (Stage B)

The core constructs of the integrated Promoting Action on Research Implementation in Health Services [24] framework were considered to facilitate innovation in the development of the dashboard: (1) innovation refers to the alignment of new insights including the predictive risk of falls captured by the dashboard with the current priorities and practices for falls prevention at RACFs, (2) recipients refer to the RACFs staff and other external stakeholders who will be affected by and influence the implementation of the dashboard, (3) context refers to the RACFs including the organization (inner layer) and wider health system (outer layer) around falls prevention and management, and (4) facilitation refers to the process that activates the implementation through assessing the characteristics of the three previously mentioned constructs.

The core constructs were integrated into the dashboard development process and the design was meticulously refined through a series of workshops (ie, facilitation) involving diverse teams of experts from various layers. These teams included general practitioners, geriatricians, allied health staff from RACFs (including nurses and clinical care managers), and human-computer interaction experts (ie, recipients). The collaborative workshops aimed to identify and implement enhancements to the dashboard, ensuring that it caters to the needs and preferences of all relevant stakeholders (ie, innovation). The design features (eg, color, graph types, positioning, filters, slices, and data views) and content for decision-making (ie, information included or missing on the dashboard and predictive model output) were considered. In these workshops, the development stage prototypes were presented to gather information on decision-making on client care and to identify the dashboard views that required improvements in managing falls. Current priorities and practices involved in managing falls at RACFs (ie, context) were also identified.

Descriptive and Diagnostic Analytics: Measures, Indicators, and Statistical Tests

All falls, injurious falls, and falls requiring hospitalization were identified as the main indicators in the dashboard. The total number of falls, the crude incident rate (ie, falls rate per 1000 resident days), the percentage of injurious falls, and the percentage of hospitalization were calculated for each outcome measure to enhance the decision support capability at an organizational level.

Funnel plots were used to visualize crude incident rates with 95% and 99% control limits, relative to the number of admissions. A time series plot was used to visualize the percentage of injurious falls and hospitalization due to falls.

Predictive Analytics With Resident Tailored Recommendations: Dynamic Falls Predictive and Monitoring Model

A dynamic predictive tool was developed to classify each resident's risk of falls using a technique called stratified landmarking and reported in detail elsewhere [17]. Initially, 116 variables were screened, and the most relevant variables on the fall outcome were selected using Collett's [25] variable selection approach. Two separate models based on dementia status were developed [17]. The final two models were based on patient demographics, dementia status, fall risk–increasing comorbidities such as cerebrovascular accident, visual impairment, osteoporosis or fracture, medication use, falls history, other risk factors such as psychological status, mobility status, behavior status, and activities of daily living [17].

The predictive model used in this study was developed and internally validated using retrospective data from the same provider [17]. The scores based on model coefficients and baseline hazards were then embedded into the dashboard backend and calculated the resident's probability of falling in a near real-time manner. The predictions generated by the model are updated every 24 hours given the modifiable risk factors such as medication use, previous fall incidents, and comorbidities.

As mentioned in the introduction, relying solely on a predictive model will not improve decision-making for developing a personalized care strategy for fall prevention and management. Additionally, the dashboard that visualizes risk factors at both the resident and facility levels should be used to enhance the interpretability of these predictions and other associated risk factors of falls. This dashboard visualizes the risk factors for falls, along with changes in risk and risk stages derived from

the predictive model. Furthermore, to offer a more prescriptive approach based on these predictions, evidence-based, resident-tailored recommendations have been developed as follows.

Resident-tailored recommendations are generated using a rule-based system in the dashboard's backend. The system identifies residents' mobility issues from PH-FRAT data and detects the use of fall risk–increasing drugs (FRIDs) from daily medication records. Fall incident data are used to determine the most frequent fall locations, while a predictive model assesses the current fall risk. Based on these factors, rule-based recommendations [26-28] are created and displayed in the resident fall view (Table S4 in Multimedia Appendix 1). Additionally, strategies identified through clinical judgment in the PH-FRATs are also displayed to enhance the personalization.

Results

Descriptive Profile of the Aged Care Provider, Facilities, and Residents

The aged care provider managed 24 RACFs. During the period of dashboard development, these RACFs housed 3686 resident admissions from 3057 unique residents. The median facility size was 124 (IQR 86-160) residents. Of the resident admissions, 71% (n=2622) were permanent. The median length of stay was 373 (IQR 41-1266) days (Tables S1 and S2 in Multimedia Appendix 1).

Dashboard Evolution Within the Design Stage

The Stage A interviews and Stage B workshops identified themes and findings related to content, design, functionality, and decision support. This led the dashboard to evolve through a series of iterations (Multimedia Appendix 2). Two dashboard views namely, organizational falls and facility falls were designed for organizational managers, while another two views were designed for care staff and focused on residents, resident falls, and resident-level medications in detail. The intervention recommendations gathered on fall prevention and management from published randomized controlled trials, government reports, systematic reviews, and meta-analyses [26-28] are included in the recommendation panel for education purposes.

Measures Developed in Dashboard Views

Textbox 1 presents the measures developed for each dashboard view. Given the interrelatedness of these views, certain measures used at the organizational level (such as total falls and falls rate [per 1000 resident days]) were also included at the facility level, while certain measures used at the resident level (ie, FRIDs use in last three days and polypharmacy indicator) were included in the medications in detail view.

Textbox 1. Measures available in the dashboard.

Organizational falls

- Total falls
- Total falls requiring hospitalization
- Total injurious falls
- Falls rate (per 1000 resident days)
- Falls requiring hospitalization rate (per 1000 resident days)
- Injurious falls rate (per 1000 resident days)
- Number of admissions per selected period
- Resident days per selected period

Resident falls

- Current falls risk
- Current risk category
- Daily percentage change in risk
- Falls in last 6 months
- Fall risk–increasing drugs (FRIDs) use in the last three days
- Change in FRIDs use in last three days
- Polypharmacy indicator
- Antipsychotic indicator
- Recent falls risk assessment tool (FRAT) assessment date
- Recent cognitive status identified by Peninsula Health-FRAT
- Recent psychological status

Facility falls

- Total number of residents
- Residents having stage 1 pressure injury
- Residents having stage 2 pressure injury
- Residents having stage 3 pressure injury
- Residents having stage 4 pressure injury
- Residents having unstageable pressure injuries
- Residents having deep tissue pressure injuries
- Residents with a pressure injury

Medications in detail

- Polypharmacy indicator – without pro re nata and short course
- Daily FRIDs indicator
- Daily change in FRIDs
- Daily change in opioid use
- Daily change in anxiolytics
- Daily change in hypnotic and sedatives
- Daily change in beta blocking agents
- Daily change in vasodilators
- Daily change in antidepressants

Dashboard Architecture

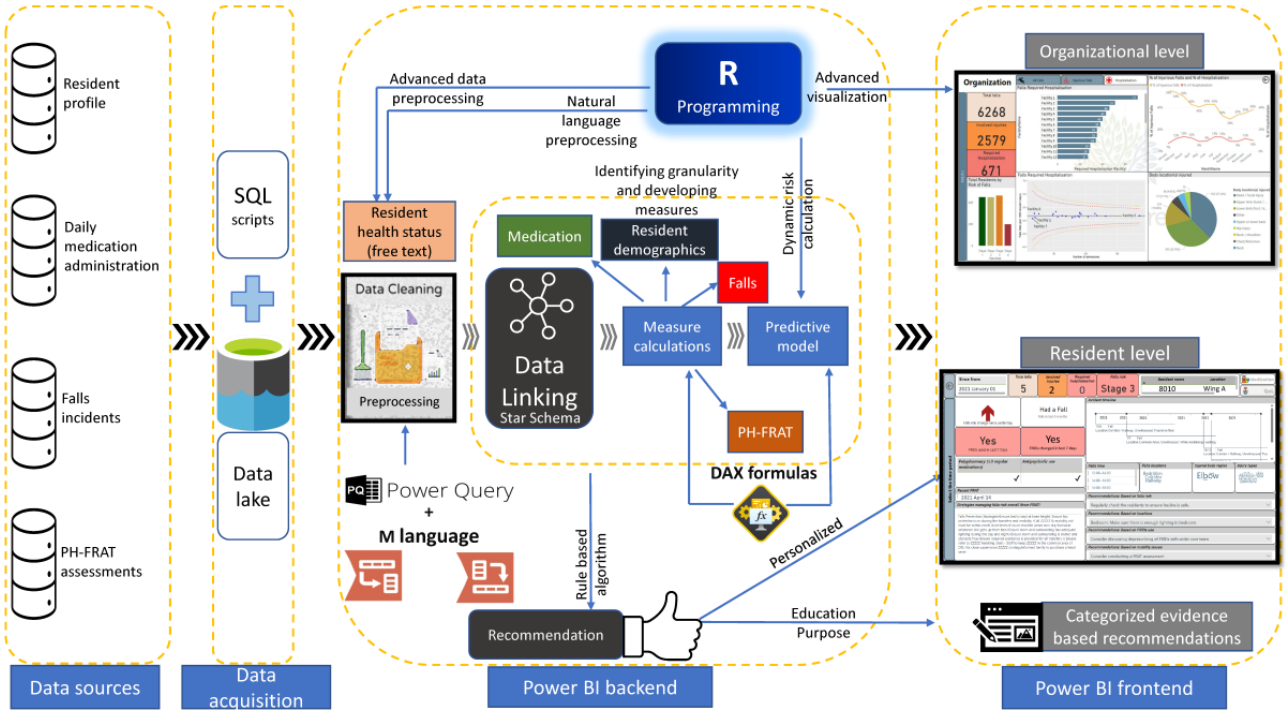
Data flows through 4 layers, to formulate the necessary information (Figure 3).

The first layer includes the data sources where the aged care provider collects and stores data. The next layer is the data acquisition layer. It includes data extraction scripts such as SQL scripts, which collect data from the previous layer. As the dashboard was developed without a direct connection to the above data silos (sources), the data extractions at the beginning were done using SQL scripts for each data silo, including the original variable names and formats. This layer is also helpful in maintaining the reliability of the dashboard when underlying IT systems change by storing the original form of data that is compatible with the dashboard backend, which is the next layer. At implementation, this layer will be replaced with a data lake that feeds provider-level data to the dashboard backend.

Then the data extracted from the data lake are preprocessed and linked at the backend of the dashboard using the data model schema discussed (Figure 2). Previously identified measures and indicators are developed in this layer, to support the visualization of data at the front end of the dashboard. For this purpose, we used power query M language, DAX, and R programming. Risk calculations of the prediction model were carried out in this stage using R programming.

In the fourth layer, the information requirement of stakeholders is satisfied by visualization and reporting. The front end of the dashboard includes five views namely organizational falls, facility falls, resident falls, medications in detail, and recommendation panel. The filters are applied to various dashboard views depending on user needs as identified during the requirement elicitation and dashboard design stages. Moreover, statistical plots are used to enhance valid comparisons of indicators between facilities and to support decision-making.

Figure 3. The dashboard architecture with four layers. DAX: Data Analysis Expression; PH-FRAT: Peninsula Health falls risk assessment tool.



Dashboard View: Organizational Falls and Facility Falls

The organizational falls and facility falls are accessible by organizational and facility managers. Please note that some information about these views has been removed or modified to deidentify the residents and facility details.

The organizational falls view (Figure 4) includes organizational and facility-level information relevant to fall incidence. The time period and facility names are included as the filters. The information on this view can be updated with the values selected for these filters. The total number of falls along with injuries, and falls that required hospitalization are reported at the organizational level. The funnel plot in the view visualizes the crude fall rate. Users can visualize this plot for the three outcome measures namely all falls, injurious falls, and falls-related hospitalization. Funnel plots with the crude incident rates for each outcome measure enable valid comparisons between the facilities, as it can identify the outlying facilities with higher fall rates related to the number of admissions during the study period. The time series plot with a percentage of injurious falls and a percentage of hospitalization enables the user to investigate overall patterns of falls occurring in selected facilities or at an organizational level. This feature is helpful in identifying the longitudinal trend of injurious falls and hospitalization due to falls in facilities. The time series plot with the most frequent time of falls enables the user to investigate the time pattern of fall incidents. The most frequent locations of falls are linked with the time of falls by a Sankey plot. These plots are dynamic and responsive to the facility selected from the filters. Furthermore, the table with fall risk stages indicates the number of residents within each risk classification predicted by the

embedded predictive model. The most frequent injured body locations are also reported in a table.

The facility falls view (Figure 5) provides more details of the facility-specific falls along with facility characteristics. This view can also be accessed through the organizational level view using the "drill through" functionality in PBI. This view highlights the facility-specific fall rate per 1000 resident days for a specified time. To enhance the interpretability of the predictive model, resident-specific risk factors extracted from the model are reported in this view. Importantly, residents classified under each risk classification along with changes in fall risk compared to the previous day can be accessed in this view for effective fall prevention and management. The view also identifies comorbidities that increase fall risk (eg, dementia) and modifiable risk factors (eg, recent falls within the last six months, frequent locations and times of falls, mobility issues, psychological status, and the use of antipsychotic, opioid, and analgesic medications), all color-coded based on the level of risk identified by the predictive model. Additionally, facility-specific monthly indicators including four national mandatory quality indicators namely the percentage of care recipients: who experienced one or more falls; who were prescribed nine or more medications; who received antipsychotic medications; and with pressure injuries are also reported in this view. This view will also be included in a daily email notification to users once the dashboard is implemented.

Figure 4. Organizational falls view.

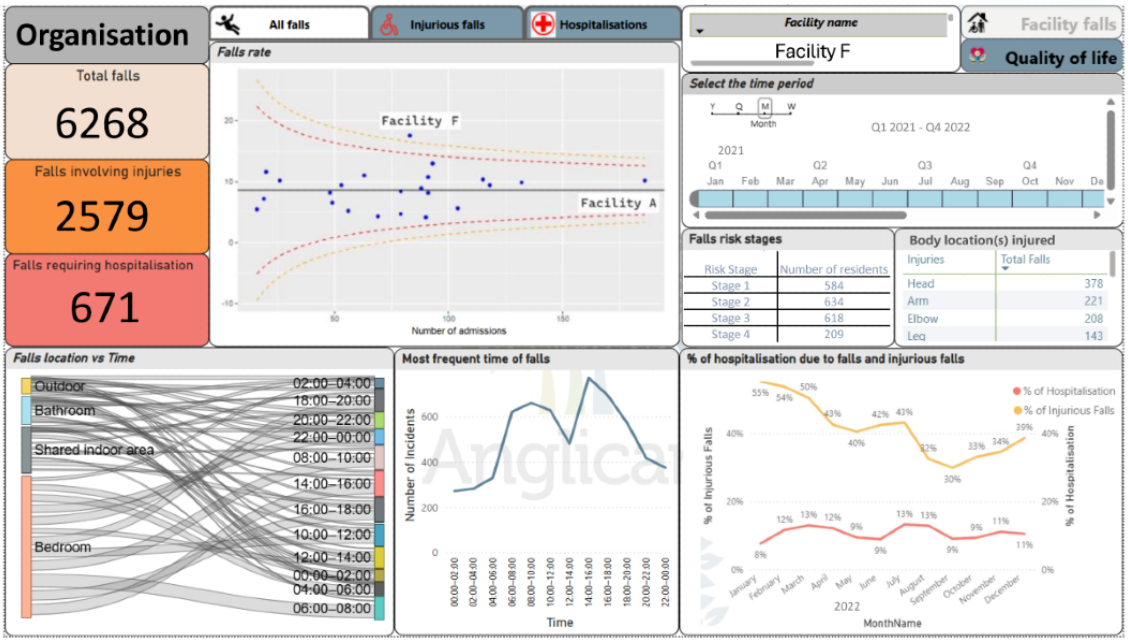
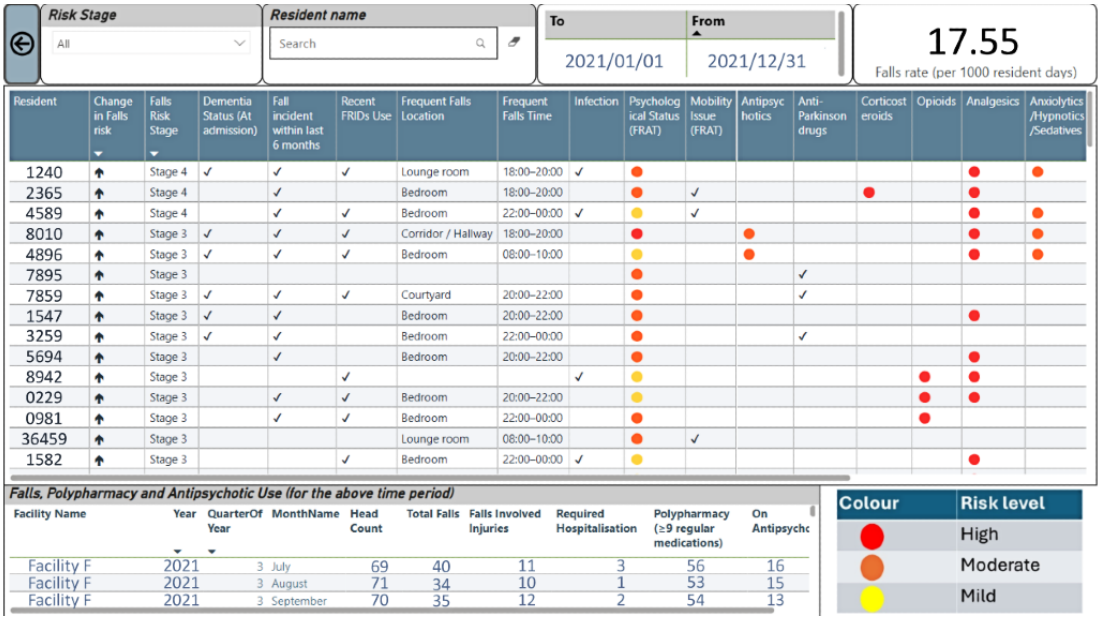


Figure 5. Facility falls view. FRAT: falls risk assessment tool.

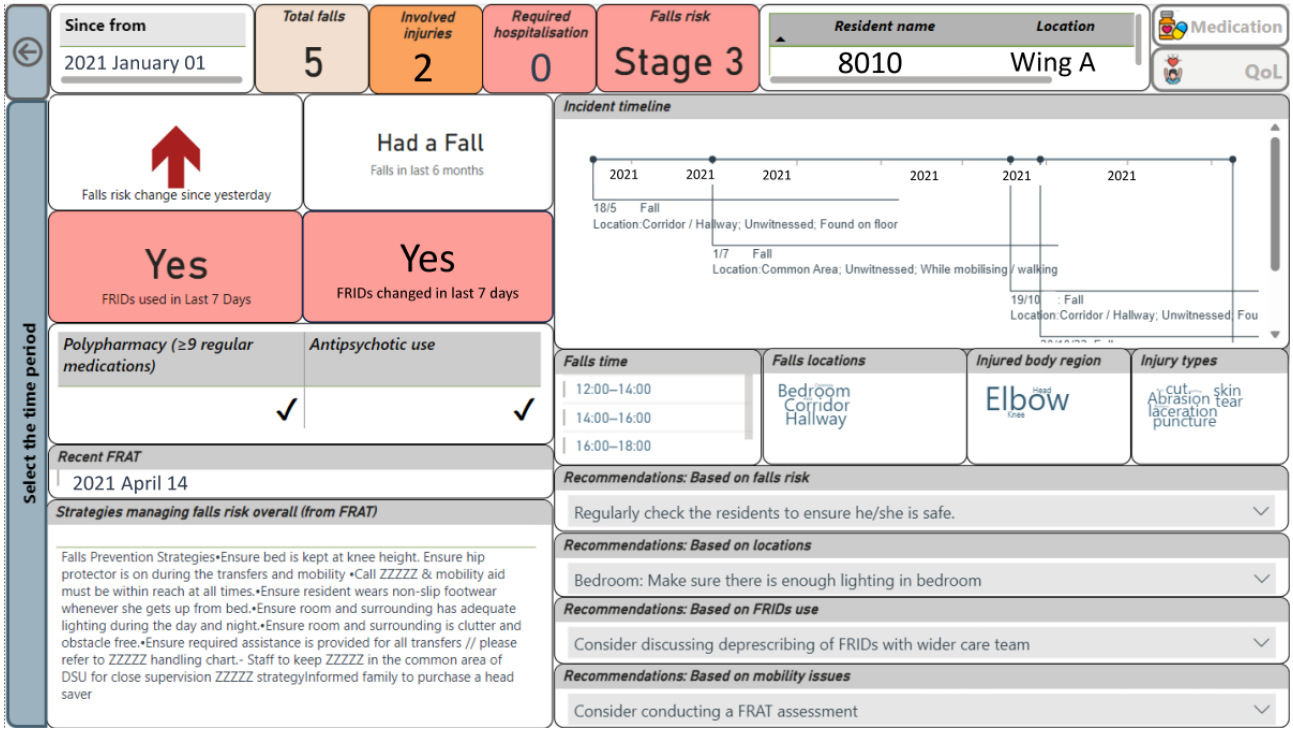


Dashboard View: Resident Falls and the Detailed Medications

Care staff will have access to the resident falls view, which includes the time period, facility name, falls risk classification (ie, derived from the predictive model), and resident identifier as filters.

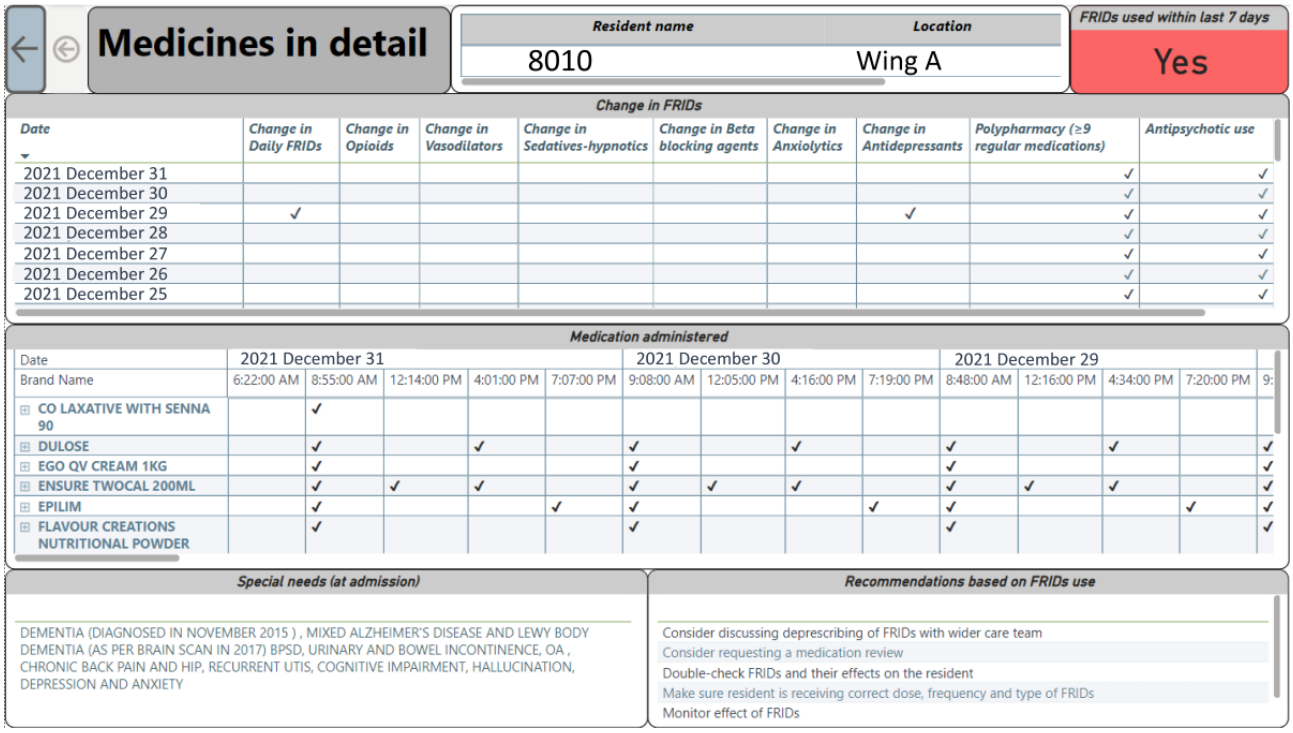
The resident falls view focuses on the resident-specific falls information along with recommendations to prevent falls (Figure 6). This can also be accessed using the facility falls view using drill-through functionality in PBI. The polypharmacy, FRIDs, and antipsychotic medication use are identified using the resident's daily medication administration data. This view presents the current risk category along with the total number of falls faced by the resident during their stay. Risk classification derived from the predictive model is shown using color codes (ie, green to red) and text (ie, stages 1 to 4). Most importantly, the number of falls faced within the last 6 months is shown in this view. Patient-specific fall history is also visualized using a timeline along with the details of the events (ie, event type, injury status, and witnessed or unwitnessed). The sequence and the gaps between the fall incidents are highlighted in this plot. Furthermore, the word clouds visualize the injured body regions, injury types, locations, and time of the falls. At the bottom of the view, personalized recommendations are provided to the resident, which aligns with falls risk categorization, falls history, medication use, and mobility issues.

Figure 6. Resident falls view. FRAT: falls risk assessment tool; FRID: fall risk–increasing drug.



The medications in detail view (Figure 7) focuses on the resident-specific daily medication use. This view can be assessed using the "drill through" functionality enabled in the resident falls view. Here, changes in FRIDs (ie, whether the resident has previously used the drug, whether a medication was omitted, or a new drug added to the medication profile) within the last three days are identified and displayed, along with polypharmacy and antipsychotic medication indicators. The resident's use of FRIDs is highlighted at the top. Daily medication administration records can also be accessed using this view. The special needs (ie, comorbidities and allergies) identified at the admission are included in the view to highlight the comorbidities, special needs, or risk factors of the resident.

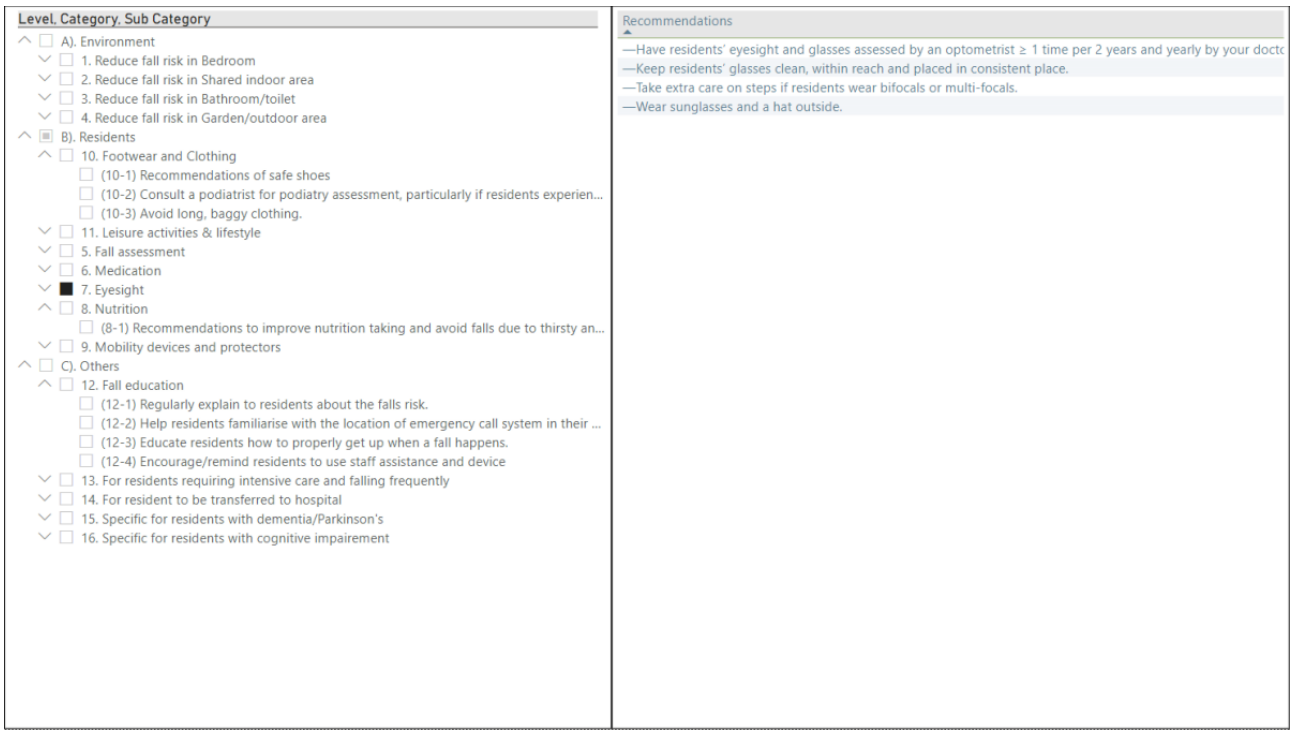
Figure 7. Medicines in detail. FRID: fall risk–increasing drug.



Dashboard View: Recommendation Panel

As falls are affected by intrinsic and extrinsic factors [29], recommendations are grouped and reported in the recommendations panel (Figure 8) in three groups namely: resident (intrinsic), environment (extrinsic), and others followed by subcategorizations.

Figure 8. Recommendation panel.



Illustrative Use Scenario

Organizational Managers

A total of 6268 fall incidents were reported from the 3057 unique residents (729,410 resident days; Table S3 in Multimedia Appendix 1) between January 1, 2021, and December 31, 2021 (Figure 4). Of these, 2579 (41%) falls were injurious and 671 (11%) falls required hospitalization. The average fall rate was 8.36 per 1000 resident days for the study period (funnel plot). The percentage of injurious falls shows an overall gradual decline from 56% to 38%, whereas the percentage of falls requiring hospitalization is stable between 7% and 13% between

January 1, 2021, and December 31, 2021. The majority of falls occurred in residents' rooms. Injurious falls most often involved injuries to the head and face followed by upper limbs. When considering all falls, Facility A reported the highest number of falls. Facility F had the sixth-highest number of falls. However, the rate was above the 99% upper control limit (funnel plot), indicating a higher fall rate (per 1000 resident days) compared to other facilities with a similar number of admissions. Therefore, organizational managers can prioritize Facility F in intervening. They also can select Facility F and drill through to the facility falls view for a more detailed investigation.

Facility Managers

A total of 83 admissions were reported for facility F during the study period, with a fall rate of 17.55 per 1000 resident days (Figure 5). As of December 31, 2021, 3 (4%) residents were categorized with a fall risk of stage 4, while 13 (18%) residents were categorized with a risk of stage 3. Most of these residents had experienced a fall within the last six months and were on FRIDs. The resident bedroom was identified as the most frequent location for falls among the majority of residents. The majority of residents in this facility had cognitive impairment and dementia. It was clear that 54 (77%) and 13 (18%) residents were on polypharmacy and antipsychotics, respectively, for September 2021. Organizational managers and facility managers can use the dashboard information to identify facility-level risk factors and assist them with mandatory quality indicator reporting.

Nurse or Personal Carer

On the facility falls view, a stage 3 resident (ie, resident: 8010) was selected and drilled down to the resident falls view to analyze the resident-level risk factors, fall incident characteristics, and personalized recommendations to prevent and manage future fall risk (Figures 5 and 6). Risk factors: resident 8010 was in stage 3 (as of December 31, 2021) with delirium, dementia, and stroke (Figure 7). The resident was on polypharmacy and antipsychotic medications as of December 30, 2021. There was an increase in fall risk compared to December 30, 2021, which was generated by the embedded predictive model. The resident had five falls during the study period, of which, two were injurious. When considering the resident's incidents timeline, four fall incidents had occurred in the last six months. The resident was on polypharmacy along with FRIDs (as of December 31, 2021) where he had a change in FRIDs in the last three days (ie, December 29-31, 2021). Characteristics of fall incidents: most of these incidents occurred during the afternoon in the resident's bedroom including the hallway and caused injuries to the elbow, head, and knee. The most recent FRAT assessment was conducted in April 2021. Recommendations to prevent future falls: according to the recent FRAT, the identified strategies in managing fall risk were displayed along with prescriptive recommendations such as "regular monitoring," "assessing environmental risk in toilet or bathroom and bedroom," and "assessing FRIDs use."

For a further detailed view of medication use, the resident's medications in detail view (Figure 7) can be assessed. Moreover, carers can use the information provided in the recommendation panel when assessing the fall risk for the resident, based on the resident-level risk factors (eg, dementia), environmental risk factors (ie, bedroom), etc.

Discussion

Principal Findings

Our dashboard adapted and refined the dashboard development process outlined by Staron [10], with an emphasis on understanding the specific domain of residential aged care and their current work practices around fall prevention and management. This was achieved through continuous stakeholder discussions (ie, Stages A and B), quantitative data analysis on the extracted datasets, and reviewing the published literature. This work produces important new resources for the current state of falls prevention in RACFs along with the extent of evidence-based prevention strategies. The dashboard design stage was an iterative process that required close collaboration with the users and its stakeholders (ie, Stage B). This was the most important stage of the dashboard development as it identified the data presentation for dashboard views along with the quantitative measures to be included for fall prevention. To enhance the technology implementation process within the context of fall prevention and management, the four constructs of the integrated Promoting Action on Research Implementation in Health Services framework were followed throughout the dashboard development process.

The core innovation of this study comes from several novel aspects. The Royal Commission into Aged Care Quality and Safety has noted that the residential aged care sector has lagged behind other health care sectors in adopting and implementing technology [30]. It is also found that information systems and processes in these settings are both underdeveloped and inadequately integrated [31]. Furthermore, successfully connecting multiple disparate data sources in aged care is uncommon. This may be due to various factors, such as data being collected by different systems and third-party software vendors being inflexible in incorporating unique identifiers (ie, resident identifiers) across systems. Therefore, this study emphasizes the value of linking existing datasets and visualizing the linked information to support falls prevention. The resident falls view in the dashboard presents the linked data from four data sources and supports to identify resident-level risk factors, fall incident characteristics, and personalized recommendations for fall prevention. This allows users to access crucial information at a glance, rather than navigating across different systems, thus demonstrating significant added value in preventing falls.

The dashboard provides staff with easy access to synthesized, evidence-based recommendations for effective fall prevention strategies. In the dashboard, external evidence on fall prevention and management (eg, published prevention recommendations) has been aligned with current priorities and practices at RACFs (such as PH-FRATs). Current practices, such as the fall prevention strategies identified through PH-FRATs, were embedded, along with rule-based recommendations. Additionally, a separate recommendation panel was established to support decision-making around fall risk assessments.

Another novel aspect of the dashboard is the embedded predictive model. Predicting fall risk is challenging because of the numerous contributing factors [32,33], including medical conditions, medications, functional status, behavior, physiological aspects, and environmental conditions. Many of these factors, such as medication use, can vary over time [34]. The dynamic predictive model embedded in the dashboard is therefore an important part of the dashboard. It uses near real-time data and helps to classify the resident risk of falls on a daily basis. The incorporation of the predictive algorithm in the dashboard goes beyond what has been possible with existing fall risk assessments as it dynamically predicts the risk and links it to residents' fall risk factors and personalized strategies. This provides timely and meaningful risk prediction along with evidence-based recommendations based on resident profile, to the caring staff to facilitate person-centered fall prevention and improve outcomes.

Finally, the dashboard is innovative because it adopts a holistic approach to managing falls in aged care residents by integrating multiple data sources and making this information easily accessible to different stakeholders including managers and carers. The granular resident-level data are aggregated to the facility and organizational levels, allowing users to navigate seamlessly between these levels with just a few clicks. This structure enables users to drill down from the organizational level to individual residents or vice versa, providing a comprehensive and detailed view of fall risk and prevention strategies.

Challenges Faced and Strategies to Overcome

There were many challenges faced when developing this dashboard. Initially, the dashboard was developed using datasets that were extracted from the provider rather than directly connecting to the database. This leads to the integration process becoming more challenging unless the variables are in standard format as in the database. Moreover, our partner has changed its incident reporting system and daily medication management system within the development process. More robust processes were required to be implemented on the backend of the dashboard to overcome these challenges. The data acquisition layer discussed in dashboard architecture was important in this aspect. This layer acts as a data reservoir. Here, the initial data are extracted from original sources, stored in the original format, and then transformed into the data tables that the dashboard requires. This layer enabled the dashboard to be sustainable in the long run amid the underlying system changes.

The data model in the backend of the dashboard also plays a critical role in facilitating the above process. This is crucial as the measures were developed using the linkages between the tables. Adapting a star schema as much as possible within the dashboard backend has promoted several advantages within the study. Simplicity and ease of use of the data model during various stages of the development process can be noted. Visualization and embedding of analytical measures within the dashboard were simplified without complex data integration. This has streamlined the dashboard development during the co-design stage while accelerating the time of incorporating various views on fall prevention and management. Flexibility and scalability can be noted as another advantage where the schema was helpful in adapting the dashboard design to changing RACF requirements in managing falls while incorporating additional data fields and sources with minimum modifications to the existing structure. Therefore, overall, adapting the star schema as much as possible within the dashboard development process was an excellent choice due to its ability to incorporate changing user requirements in the context of fall management at RACFs.

Furthermore, for the daily calculation of changes in fall risk, this study uses R programming. This decision stems from PBI's inability to retain and save daily risk calculations. To address this limitation, the designers have integrated R programming code into the Power Query Editor, enabling the storage of these risk calculations on a shared drive. The code automatically saves these calculations on a daily basis during the data refresh cycle, occurring every 24 hours. Subsequently, on a daily basis, these stored files (ie, previous-day risk calculations) are retrieved into the PBI backend to calculate the difference between current and previous risks at the resident level.

The integration of R programming into the dashboard backend has introduced a technical skills gap between the partner organization and the dashboard designers. Consequently, workshops were organized with the partner organization's IT teams to provide training on R programming, bridging the gap in statistical analysis using R programming. Additionally, line-by-line code interpretation was incorporated into the embedded codes to facilitate a seamless transition during dashboard implementation.

Implications for Policy, Practice, and Future Research

This study highlights the potential of using a fall prevention and management dashboard to enhance resident outcomes in aged care through the effective use of information systems. Health dashboards facilitate decision-making and are required to be used in conjunction with clinical judgment, staff education, training, and individualized tailored responses. Future policy and practice must ensure the interoperability of information systems in aged care, involve key stakeholders in development and implementation, and allocate appropriate funding for these processes. Future research should also focus on expanding the fall prediction dashboard and incorporating other care needs in residential aged-care facilities, such as hospitalization and wound management.

Conclusions

In this study, a comprehensive process for developing a fall prevention and management dashboard with descriptive, predictive, and prescriptive analytics for RACFs was thoroughly discussed along with a sustainable architecture. The key focus was to provide a holistic approach to managing falls by linking existing data sources at the provider level. This highly applied and translational research demonstrated how EHRs can be used along with technology to deliver evidence-based information to aged care allied health workers and managers, to drive better aged care quality and safety.

Acknowledgments

We thank our partners and collaborators including Anglicare, Northern Sydney Local Health District, Sydney North Primary Health Network, the Deeble Institute for Health Policy Research, and the Australian Aged Care Quality and Safety Commission.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the privacy of the individuals who participated in the study but are available from the corresponding author upon reasonable request.

Authors' Contributions

The dashboard was technically designed by SSMS. The co-design process of the dashboard involved SSMS, NW, ADN, KS, LD, IM, CIM, and GH. The manuscript was drafted by SSMS and revised by NW, ADN, KS, GH, and IM. Both the manuscript draft and system were reviewed by JIW.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 28 KB - aging_v8i1e63609_app1.docx]

Multimedia Appendix 2
Dashboard evolution within the design stage.
[PNG File , 534 KB - aging_v8i1e63609_app2.png]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2022-11-25]
2. Falls and fractures: effective interventions in health and social care. Department of Health. 2009. URL: https://laterlifetraining.co.uk/wp-content/uploads/2011/12/FF_Effective-Interventions-in-health-and-social-care.pdf [accessed 2024-05-05]
3. Ward JA, Harden M, Gibson RE, Byles JE. A cluster randomised controlled trial to prevent injury due to falls in a residential aged care population. Med J Aust 2010;192(6):319-322. [doi: 10.5694/j.1326-5377.2010.tb03530.x] [Medline: 20230348]
4. Waldron N, Hill AM, Barker A. Falls prevention in older adults—assessment and management. Aust Fam Physician 2012 Dec;41(12):930-935. [doi: 10.3316/informit.998232410954897] [Medline: 23210114]
5. Cameron ID, Dyer SM, Panagoda CE, Murray GR, Hill KD, Cumming RG, et al. Interventions for preventing falls in older people in care facilities and hospitals. Cochrane Database Syst Rev 2018;9(9):CD005465 [FREE Full text] [doi: 10.1002/14651858.CD005465.pub4] [Medline: 30191554]
6. Wabe N, Siette J, Seaman KL, Nguyen AD, Raban MZ, Close JCT, et al. The use and predictive performance of the Peninsula Health Falls Risk Assessment Tool (PH-FRAT) in 25 residential aged care facilities: a retrospective cohort study using routinely collected data. BMC Geriatr 2022;22(1):271 [FREE Full text] [doi: 10.1186/s12877-022-02973-0] [Medline: 35365078]
7. Park SH. Tools for assessing fall risk in the elderly: a systematic review and meta-analysis. Aging Clin Exp Res 2018;30(1):1-16. [doi: 10.1007/s40520-017-0749-0] [Medline: 28374345]
8. Ludlow K, Westbrook J, Jorgensen M, Lind KE, Baysari MT, Gray LC, et al. Co-designing a dashboard of predictive analytics and decision support to drive care quality and client outcomes in aged care: a mixed-method study protocol. BMJ Open 2021 Aug 25;11(8):e048657 [FREE Full text] [doi: 10.1136/bmjopen-2021-048657] [Medline: 34433599]
9. Thapa R, Garikipati A, Shokouhi S, Hurtado M, Barnes G, Hoffman J, et al. Predicting falls in long-term care facilities: machine learning study. JMIR Aging 2022 Apr 01;5(2):e35373 [FREE Full text] [doi: 10.2196/35373] [Medline: 35363146]
10. Staron M. Dashboard development guide—how to build sustainable and useful dashboards to support software development and maintenance. 2015. URL: https://core.ac.uk/download/pdf/43560952.pdf [accessed 2025-03-08]
11. What is the network of patient safety databases? Agency for Healthcare Research and Quality. 2023. URL: https://www.ahrq.gov/npsd/what-is-npsd/index.html [accessed 2025-03-08]
12. Live Stronger. 2023. URL: https://www.livestronger.org.nz/assets/Uploads/falls-and-fractures-indicator-methodology.pdf [accessed 2025-03-08]
13. Hjørland B, Albrechtsen H. Toward a new horizon in information science: domain-analysis. J Am Soc Inf Sci 1995;46(6):400-425. [doi: 10.1002/(sici)1097-4571(199507)46:6<400::aid-asi2>3.0.co;2-y]

14. Falls risk assessment tool (FRAT) instructions. Department of Health & Human Services. 2016. URL: https://training.aacs.com.au/wp-content/uploads/2016/08/Falls-Risk-Assessment-Tool-FRAT.pdf [accessed 2024-05-01]
15. Meulenbroeks I, Mercado C, Gates P, Nguyen A, Seaman K, Wabe N, et al. Effectiveness of fall prevention interventions in residential aged care and community settings: an umbrella review. BMC Geriatr 2024;24(1):75 [FREE Full text] [doi: 10.1186/s12877-023-04624-4] [Medline: 38243175]
16. Seaman K, Ludlow K, Wabe N, Dodds L, Siette J, Nguyen A, et al. The use of predictive fall models for older adults receiving aged care, using routinely collected electronic health record data: a systematic review. BMC Geriatr 2022;22(1):210 [FREE Full text] [doi: 10.1186/s12877-022-02901-2] [Medline: 35291948]
17. Wabe N, Meulenbroeks I, Huang G, Silva S, Gray L, Close J, et al. Development and internal validation of a dynamic fall risk prediction and monitoring tool in aged care using routinely collected electronic health data: a landmarking approach. J Am Med Inform Assoc 2024;31(5):1113-1125 [FREE Full text] [doi: 10.1093/jamia/ocae058] [Medline: 38531675]
18. Pacheco C, Garcia I, Calvo-Manzano JA, Reyes M. Measuring and improving software requirements elicitation in a small-sized software organization: a lightweight implementation of ISO/IEC/IEEE 15939:2017—systems and software engineering—measurement process. Requir Eng 2022;28(2):257-281. [doi: 10.1007/s00766-022-00394-4]
19. Lind KE, Raban MZ, Brett L, Jorgensen ML, Georgiou A, Westbrook JI. Measuring the prevalence of 60 health conditions in older Australians in residential aged care with electronic health records: a retrospective dynamic cohort study. Popul Health Metr 2020;18(1):25 [FREE Full text] [doi: 10.1186/s12963-020-00234-z] [Medline: 33032628]
20. Aged care health status algorithms. Macquarie University, Australia. 2020. URL: https://www.mq.edu.au/research/research-centres-groups-and-facilities/healthy-people/centres/australian-institute-of-health-innovation/Research-Streams/the-aged-care-evaluation-and-research-acer-team/aged-care-health-status-algorithm [accessed 2023-12-01]
21. Anatomical therapeutic chemical (ATC) classification. World Health Organization. URL: https://www.who.int/tools/atc-ddd-toolkit/atc-classification [accessed 2022-01-05]
22. Kimball R, Ross M. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling. Indianapolis, IN: Wiley; 2013.
23. Deckler G, Powell B, Gordon L. Mastering Microsoft Power BI: Expert Techniques to Create Interactive Insights for Effective Data Analytics and Business Intelligence. Birmingham, England: Packt; 2022.
24. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implement Sci 2016 Mar 10;11:33 [FREE Full text] [doi: 10.1186/s13012-016-0398-2] [Medline: 27013464]
25. Collett D. Modelling Survival Data in Medical Research. Boca Raton, USA: CRC Press LLC; 2003.
26. Vieira ER, Palmer RC, Chaves PHM. Prevention of falls in older people living in the community. BMJ 2016;353:i1419. [doi: 10.1136/bmj.i1419] [Medline: 27125497]
27. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022;51(9):afac205 [FREE Full text] [doi: 10.1093/ageing/afac205] [Medline: 36178003]
28. Don't fall for it, falls can be prevented!: a guide to preventing falls for older people. Australian Government. 2011. URL: https://www.health.gov.au/sites/default/files/documents/2021/04/don-t-fall-for-it-falls-can-be-prevented.pdf [accessed 2025-03-08]
29. Statewide Action Plan: Falls Prevention in Older People 2002-2006. Brisbane: Queensland Health; 2002.
30. Final report: care, dignity and respect. Royal Commission into Aged Care Quality and Safet. 2021. URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2025-03-08]
31. Davis J, Morgans A, Burgess S. Information management in the Australian aged care setting. Health Inf Manage J 2017;46(1):3-14. [doi: 10.1177/1833358316639434] [Medline: 27105477]
32. Graafmans WC, Ooms ME, Hofstee HM, Bezemer PD, Bouter LM, Lips P. Falls in the elderly: a prospective study of risk factors and risk profiles. Am J Epidemiol 1996;143(11):1129-1136. [doi: 10.1093/oxfordjournals.aje.a008690] [Medline: 8633602]
33. Klenk J, Becker C, Palumbo P, Schwickert L, Rapp K, Helbostad JL, et al. Conceptualizing a dynamic fall risk model including intrinsic risks and exposures. J Am Med Dir Assoc 2017;18(11):921-927. [doi: 10.1016/j.jamda.2017.08.001] [Medline: 28916290]
34. Rajagopalan R, Litvan I, Jung T. Fall prediction and prevention systems: recent trends, challenges, and future research directions. Sensors (Basel) 2017 Nov 01;17(11):2509 [FREE Full text] [doi: 10.3390/s17112509] [Medline: 29104256]

Abbreviations

DAX: Data Analysis Expression
EHR: electronic health record
FRAT: falls risk assessment tool
FRID: fall risk–increasing drug
PBI: Power BI

PH-FRAT: Peninsula Health Falls Risk Assessment Tool
RACF: residential aged care facility

Edited by Y Jiang; submitted 26.06.24; peer-reviewed by J Bhowmik, J Walsh; comments to author 20.09.24; revised version received 10.10.24; accepted 28.02.25; published 07.04.25.

Please cite as:
Silva SSM, Wabe N, Nguyen AD, Seaman K, Huang G, Dodds L, Meulenbroeks I, Mercado CI, Westbrook JI
Development of a Predictive Dashboard With Prescriptive Decision Support for Falls Prevention in Residential Aged Care: User-Centered Design Approach
JMIR Aging 2025;8:e63609
URL: https://aging.jmir.org/2025/1/e63609
doi:10.2196/63609
PMID:

Review

Evaluating the User Experience and Usability of Game-Based Cognitive Assessments for Older People: Systematic Review

Rhys Mantell[1,2], BEc, MPH; Ye In Jane Hwang[1,2], PhD; Matthew Dark[1], BSc, MPH; Kylie Radford[2,3,4], PhD; Michael Kasumovic[5,6], PhD; Lauren Monds[7], PhD; Peter W Schofield[8,9], MSc, MD; Tony Butler[1], PhD; Adrienne Withall[2,4], PhD

[1]School of Population Health, Faculty of Medicine and Health, University of New South Wales (UNSW), Sydney, Australia

[2]Ageing Futures Institute, University of New South Wales, Sydney, Australia

[3]Neuroscience Research Australia (NeuRA), Sydney, Australia

[4]School of Psychology, Faculty of Science, University of New South Wales, Sydney, Australia

[5]School of Biological, Earth & Environmental Sciences, Faculty of Science, University of New South Wales, Sydney, Australia

[6]Arludo Holdings Pty Ltd, Sydney, Australia

[7]Specialty of Addiction Medicine, Central Clinical School, The University of Sydney, Sydney, Australia

[8]Neuropsychiatry Service, Hunter New England Health, Newcastle, Australia

[9]School of Medicine & Public Health, University of Newcastle, Newcastle, Australia

Corresponding Author:
Rhys Mantell, BEc, MPH
School of Population Health
Faculty of Medicine and Health
University of New South Wales (UNSW)
Level 2, Samuels Building, Gate 11, Botany Street
Kensington
Sydney, 2052
Australia
Phone: 61 9385 3136
Email: r.mantell@unsw.edu.au

Abstract

Background: Game-based cognitive assessments (GBCAs) have the potential to transform the field of cognitive testing by enabling more effective screening of age-related cognitive decline. However, we lack a strong understanding of the usability and overall user experience of these games. This is a risk because the primary target users for GBCAs, older people, are seldom involved in game design research and development.

Objective: This study aims to address this gap by investigating the usability, acceptability, and enjoyability of GBCAs for older people.

Methods: This study followed established practices for undertaking evidence-based systematic reviews.

Results: The initial database search returned 15,232 records. After a thorough screening process, 8 studies remained for extraction and analysis. A synthesis of the included papers identified 2 overlapping yet distinct areas of focus: system usability and subjective user experience. Usability scores were mostly positive across the studies included. However, in several of the game studies, older adults and those with cognitive impairment tended to find GBCAs less usable. This trend was observed even when the games were explicitly designed for these populations, and the tasks were simplistic and representative of basic daily activities. In our second focus area, user experience, we identified the importance of perceived challenge in mediating gameplay experience across groups. That is, generating the appropriate level of difficulty for each user is important for positive user experiences, specifically enjoyment.

Conclusions: On the basis of these findings, we identified key learnings for researchers interested in designing and developing GBCAs. These include (1) recognizing that validity is essential but not sufficient on its own; (2) clearly defining the intended user; (3) designing games that align with the unique preferences and needs of older people; and (4), whenever possible, providing each user with their optimal level of challenge.

Trial Registration: PROSPERO CRD42023433298; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023433298

(JMIR Aging 2025;8:e65252) doi:10.2196/65252

KEYWORDS

serious games; gamification; cognitive assessment; usability; user experience; aging

Introduction

Background

In recent years, there has been growing enthusiasm about the use of serious game–based technology. The primary objective of these games is to increase user engagement and motivation, thereby offering people a more pleasant environment to complete "serious" objectives that go beyond entertainment [1]. One area identified as being particularly well suited to the application of serious game technology is the field of cognitive assessment [2].

Serious games for cognitive assessment have typically been developed in 1 of 2 ways [3]. The first is "gamification," the process of building game-like features onto a cognitive task. This includes adding points, time limits, or appealing graphics to the foundational design of a traditional task. The second is to create or use an "actual" game, where the entire environment has been purposefully designed to be a game. This approach, more commonly associated with the term *serious game*, offers more creative development options, player choice, in-game competition, overarching narratives, and integrated or blended design elements. This approach is commonly based on theories of cognition or reimagining commercial games rather than reproducing a traditional cognitive task [3]. However, there remains significant confusion and inconsistency across the literature regarding what constitutes a serious game versus a gamified task. Therefore, unless otherwise noted, this paper refers to both approaches collectively as game-based cognitive assessments (GBCAs).

Traditional Cognitive Screening and the Case for Change

Traditional cognitive assessments are used by health professionals to evaluate a person's cognitive, or "thinking," functions to identify cognitive impairments or age-related declines, such as Alzheimer's disease (AD) [4,5]. Most clinicians initially use short cognitive screening tests (eg, 10-15 min) when assessing these functions. Although there are clear and well-evidenced benefits of using traditional cognitive assessment and screening tools (eg, the Montreal Cognitive Assessment has been shown to have high sensitivity [6]), there are also several issues that reduce their overall effectiveness. One concern often cited by users is that traditional screening tests can be boring and repetitive, thereby reducing user motivation and engagement during assessment. This is problematic because data obtained from individuals who are not motivated to perform a task to the best of their abilities may not be representative of their optimal performance, and this can cause misleading performance data (eg, false positives) [1]. Brief assessment tools also lack ecologic validity, meaning tasks may not translate to real-life circumstances [1,3,7]. In addition, assessments can sometimes create feelings of shame and distress [8]. Consequently, some people report significant test anxiety as well as self-stigma related to low literacy or education levels when completing cognitive tests [9,10]. Another limitation is the concerning association between cognitive performance and education level found with many screening tools [5]. Finally, some widely used neuropsychological tests are not culturally appropriate [11].

An initial attempt to address some of the concerns with traditional screening methods included digital software suites, such as CogTest and the Cambridge Neuropsychological Test Automated Battery [12]. Although these tools enable more consistent assessment and data capture, they are essentially digitized versions of traditional cognitive tests and, thus, are not gamified. Tests of this nature have encountered validation issues [12] as well as similar problems to traditional testing, such as a lack of user motivation when performing somewhat uninteresting tasks on a computer [2].

GBCAs Offer Exciting Alternatives, but They Must Be Valid, Acceptable, and Motivating

Games have recently been seen as a viable alternative to address many of these issues and improve both testing engagement and validity [1,2,12]. The benefits of using GBCAs include increased participant motivation and engagement relative to traditional tasks [2,3]. GBCAs can also offer shorter, more cost-effective, and scalable assessments [13,14]. Unlike traditional "pen-and-paper" assessments, games do not need to be initiated by a care provider and can be designed for self-administration or delivery by nonclinicians [1,14]. In addition, they may support more ecologically valid assessment through realistic context and gameplay, thereby engaging cognitive processes in ways that closely mirror real-life situations [15,16]. GBCAs can also enable consistent tracking of an individual's cognitive performance data to support the detection of subtle changes or variations in cognitive processes over time [14]. The richness of this type of data can provide invaluable population-level insight into subtle aggregate patterns, changes, or important impacts (eg, effects of exercise or smoking) on cognitive health across groups. This is an especially attractive prospect, given the increasing sophistication of machine learning models.

However, to justify their use, GBCAs must function as effective and reliable cognitive assessment tools (ie, psychometrically validated) while also being usable, acceptable, and motivating. Much of the current literature on GBCAs assumes that the simple process of gamification improves user engagement and creates a more satisfying and less tedious experience than traditional pen-and-paper or digitized neurological tests. Although many games have been found to achieve this [1-3], this is not always the case. For instance, Birk et al [17] found that adding game elements to standard cognitive tests (eg, go/no go; n-back tasks) did not improve user engagement when compared with the original task. Gamification actually reduced test performance for the go/no go task. Tong et al [18] reported difficulties in testing a carnival whack-a-mole game with older

users, also based on the go/no go paradigm, because of ergonomic issues with mobile and touchscreen devices (eg, the difficulty holding a touch device). It is also important to recognize that there may be restrictive settings or social contexts where GBCAs may be impractical, and traditional approaches may be more appropriate.

Overall, when games are not appropriately designed for their target users or contexts, they risk being neither acceptable nor motivating. Similarly, games might be skillfully designed and psychometrically validated but not adequately evaluated from users' perspectives. A systematic review of game-based interventions for neuropsychological assessment, training, and rehabilitation found that user experience evaluation was performed in <25% of interventions, with usability and enjoyability assessed in only 13% and 18%, respectively [19]. For a technology specifically aimed at increasing user engagement and enjoyment, this lack of user design input and evaluation appears to be a barrier to achieving the technology's intended outcomes.

GBCAs for Older Users

The group of interest for this study is older people (aged ≥50 years) who have or may be at risk of cognitive decline and, consequently, neurocognitive disorders (eg, dementia, such as AD). Recent research on aging has suggested that groups considered marginalized (eg, people experiencing homelessness or incarceration) are at risk of age-related conditions, such as frailty and cognitive impairment [20,21], once they reach the age of 50. Thus, to ensure adequate inclusion of people considered marginalized who are at risk of cognitive decline, we define older people as those aged ≥50 years.

Concerns about user engagement are exacerbated for game assessments aimed at detecting cognitive decline such as dementia because the target group for such tests is typically older people. This represents a challenge because games are rarely designed for or tested with older people. By contrast, there is promising—but preliminary—research to suggest that people who either have or are at risk of cognitive decline can play, attend to, and enjoy digital games, even when they do not have previous experience with touchscreen devices [22]. However, their ability to successfully complete gamified tasks appears to depend on the specific application and mechanics used within a game [22]. More broadly, the game preferences of older people seem to be different from those of younger cohorts, with studies reporting the preferences of older users gravitating toward intellectually stimulating games (ie, puzzle, educational, and strategy games) [23] that enable them to compete for high scores, require only a single player, and emphasize intellectual challenge over quick reflexes [22]. According to Blocker et al [24], games with violent content, fantasy characteristics, or interactive web-based components are not preferred by older users. A qualitative study investigating the cognitive game preferences of older people in prison reported similar findings [25], with participants suggesting a need to avoid "childish" game design, which some found condescending.

The Useful Concepts of Usability and User Experience

Across the design literature, usability and user experience are cited as major determinants of the successful adoption of any information system. Therefore, both are useful concepts when considering the acceptability, motivation, and engagement generated by any GBCA. Usability is the capacity of an object (ie, a digital game-based application) to functionally serve its intended purpose (ie, valid cognitive assessment) through qualities ranging from technical efficiency to conformity and configuration to ease of use [26]. According to the International Organization for Standardization (ISO: 9241-210:2019) [27], this refers to "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use." According to Hassenzahl and Tractinsky [28], usable systems or products enable users to achieve pragmatic objectives successfully and efficiently without obvious barriers.

On the other hand, user experience refers to the feelings a person has while interacting with a product under particular conditions [26]. This means that it goes beyond meeting the instrumental (pragmatic) need for usefulness and efficiency that good usability provides. User experience includes the need to achieve hedonistic goals [28,29], which involve "users' emotions, beliefs, preferences, perceptions, comfort, behaviors, and accomplishments that occur before, during and after use" [27]. Although there is a clear overlap, the distinction between the concepts is that good usability makes it easy for someone to use something, whereas good user experience makes using that thing feel enjoyable or satisfying.

In the context of gamified technology, usability is focused on technical features such as a seamless interface, ergonomics, clear instructions, minimal in-game bugs or errors, and other features that facilitate use, whereas user experience is concerned with impacting a person's subjective experience about the contents of a game, that is, a compelling narrative, an exciting challenge, motivation toward certain objectives, and so on [26]. There are more detailed and stratified definitions of usability and user experience across the user design literature (eg, refer to the 10 Usability Heuristics for User Interface Design by Nielsen [30]). However, consistent with our simplified definition of GBCAs, we intentionally use the basic demarcations of usability and user experience defined earlier to enable a consumable preliminary synthesis and analysis of this still-developing area of inquiry (ie, user evaluations of GBCAs for older people).

Arguably, the critical success factor for any GBCA is a coupling of user evaluation and psychometric validity. This is particularly true for older people at risk of, or presenting with, cognitive decline who are likely to have different preferences and experiences from other cohorts. However, much of the current literature is skewed toward investigating psychometric validity before or without evaluating usability and user experience. For instance, although several reviews in recent years have investigated GBCAs [2,3,31-34] and some have reported summary results on user evaluation methods (ie, how many studies conducted user evaluation) [19,35], there remains a limited understanding of the importance of user evaluation

findings. This appears to be a risk, as even the most clinically useful game in the world is unlikely to succeed at scale if it is neither functional nor engaging.

The Aims of This Study

This systematic review attempts to address some of these concerns by investigating the usability, acceptability, and enjoyability of GBCAs that have undergone psychometric validation. The intention is to better understand the nexus of clinical usefulness, usability, and user experience and, in turn, report preliminary lessons learned from user evaluations of validated GBCAs.

Methods

This study (protocol registered on PROSPERO [CRD42023433298]) followed established practices for undertaking evidence-based systematic reviews, including the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist (Multimedia Appendix 1) and the Scottish Intercollegiate Guidelines Network methodology checklist for study quality appraisal (Multimedia Appendix 2 [16,36-43]).

Data Collection, Extraction, and Quality Assessment

Independent and systematic database searches of PsycINFO, Embase, IEEE, and MEDLINE were conducted by the first author (RM) based on the predefined search terms summarized in Table 1. Broad search parameters were set, given the disparity of terminology used across the GBCA research. Relevant studies were then imported into Covidence, ready for initial screening. Title screening was manually completed by RM, where obvious exclusions and duplicates were removed. Duplicates were removed automatically via Covidence's duplicate removal function. This was followed by independent and systematic screening of relevant abstracts by 2 reviewers (RM and MD), who applied predefined inclusion and exclusion criteria to assess all abstracts.

Table 1. Database search terms. Note: the truncation symbol * (asterisk) searches for multiple variants of a word (eg, singular, plural, different conjugations, etc) all at once (eg, cogn* includes searches for cognition, cognitive, cognitively, cognizant, etc). MCI: mild cognitive impairment; VR: virtual reality.

Topic area	Search terms
Serious games	game* OR gami* OR virtual reality OR VR
	AND
Cognitive	cogn* OR MCI OR dementia OR neurocog* OR neuropsych*
	AND
Assessment	assess* OR test* OR measur* OR collect* OR detect*

Textbox 1 presents the inclusion and exclusion criteria. Important definitions and justifications relating to the inclusion and exclusion criteria are discussed more thoroughly in the Discussion of important terms subsection. When disagreements arose at the abstract screening level, the second author (YIJH) facilitated a resolution through a 3-way vote. Full-text screening was then conducted by RM and MD. Conflicts at this stage were again handled through 3-way voting consensus involving YIJH, RM, and MD. At the full-text screening stage, there were disagreements with 12 (15%) of 83 papers, which necessitated additional consensus voting. Finally, a thorough data extraction process was conducted by RM, including a risk of bias assessment. We used the Scottish Intercollegiate Guidelines Network "Methodology Checklist 4: Case Control Studies" [36] guidance to assess the risk of bias for each included study. A detailed risk of bias assessment can be found in Multimedia Appendix 2. This final extraction process was reviewed by all authors to ensure consensus and quality control.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- User evaluated
- People aged ≥50 years
- Psychometrically validated
- Assessment or screening
- Cognitive functioning
- Primary research
- 2015 onward
- Digital serious games and gamification of task

Exclusion criteria

- Did not include target population or outcome
- Intervention with high immersion (via head-mounted display, etc)
- Training or rehabilitation focused without assessment
- No psychometric validation
- No user evaluation
- No data collection performed (ie, game performance data not collected)
- Inappropriate study design (reviews, conference abstracts, etc)
- Focus on social cognition only

Discussion of Important Terms

This section summarizes and justifies several key concepts relevant to our inclusion criteria in Textbox 1, as well as our overall search strategy. It is important to emphasize that (1) our search strategy was intentionally broad to minimize the risk of missing any relevant GBCA studies and (2) our inclusion criteria were intentionally narrow to address our research aim of investigating the usability, acceptability, and enjoyability of GBCAs that have undergone psychometric validation.

User-Evaluated Study

The key criterion in this review was that a GBCA must have been user evaluated through primary research. This required the explicit inclusion of methods to measure usability, user experience, user motivation, user preference, enjoyment, or acceptability of the game.

GBCA Overview

Our definition of GBCA mirrors the broad definition included in the Introduction section. That is, both simple task gamification and more “from scratch” serious games were included. In addition, our focus was technology-based GBCAs; thus, we only included digital GBCAs (ie, games that could be played on computers, tablets, phones, etc).

Target Group

The target group for this review was older people (aged ≥50 years) who had, or were at risk of, cognitive decline and, consequently, neurocognitive disorders (eg, dementia, such as AD). Although there are numerous GBCAs that have been designed and tested to detect cognitive impairment associated with other disorders across the life course (eg, attention-deficit/hyperactivity disorder, depression, traumatic brain injury, drug- and alcohol-related cognitive deficits, etc), this paper is interested specifically in the assessment of cognitive decline in older people. This demarcation enables a manageable and specific grouping of findings.

Validated Study

To be included in this review, a GBCA must have been psychometrically validated with a sample of older users (aged ≥50 years). This validation could have occurred in the same study as the user evaluation or in a previous validation of the same game. In other words, evidence of validation was a prerequisite for inclusion in this review, but validation of the game was not necessarily required within the actual studies included in this review. According to previous definitions [3], validation included comparisons against a traditional cognitive screening tool (eg, Montreal Cognitive Assessment), a comprehensive neuropsychological battery, or a previous clinical diagnosis of cognitive impairment.

Cognitive Domains Included

A broad definition of cognitive functioning was applied for this study. Cognition refers to the brain’s ability to perceive, assimilate, organize, store, and manipulate information [44]. As such, cognitive functioning is an umbrella term that encompasses a range of integrated skills and domains that allow us to process and respond to information within our environment [44]. These domains include (but are not limited to) basic functioning (eg, processing speed and attention) and memory (eg, episodic memory, semantic memory, and prospective memory), visuospatial functions (eg, color perception and mental

rotation), and executive functions (eg, planning, and decision-making). Relevant GBCAs testing any (or all) of these domains were included in this review, with the notable exception of social cognition. The clinical assessment of social cognition and decline in older people is not as well established as these other cognitive domains and was therefore intentionally excluded. However, investigating social cognitive domains should be seen as a critical consideration for future research.

Low-Immersion Interventions

This review made a distinction between low- and high-immersion games when deciding if a study should be included. Games that were deemed highly immersive from a technological perspective, such as those requiring headset technology and immersive 3D environments, were excluded. This review focused on GBCAs in a digital format that enabled accessibility through touch or mouse click capability. The main reason for this distinction was to ensure uniformity in findings. Although highly immersive games have significant potential benefits with respect to ecologic validity and integrated assessment of multiple cognitive functions [45], they have distinctive technology requirements when compared to computer or tablet-based game systems. In addition, factors such as motion sickness [46], which can affect enjoyability, are specific to virtual 3D headset environments and would have added unwanted complexity to this synthesis. Importantly, due to the disparity in the definition of serious games, gamification, and virtual reality across the literature, some interventions have been included in this review even when the authors refer to their interventions as *3D* or *virtual reality* game applications. However, these interventions aligned with our definition of a tablet-based and low-immersion GBCA and were thus included.

2015 Onward

There is limited literature before 2015 that investigated games that were both psychometrically validated and user evaluated. There are 2 notable reviews that looked at validity before 2015 [2,34], but these reviews concluded that little work had been done in user evaluation of GBCAs before this time. In addition, a preliminary scan of literature before 2015 by the first author confirmed a lack of usability data from psychometrically validated games. Given our very broad search strategy, the significant growth of literature on GBCAs in the last 5 to 7 years, and a lack of user evaluation data before this time, a cut-off of 2015 for selection appeared justified.

Results

Overview

The initial database search on December 9, 2022, identified 15,232 records. There were 7595 articles left after removing duplicates. A total of 6946 articles were removed during the title screening for obvious exclusions, leaving 649 articles for abstract screening. Full texts were retrieved for 83 articles, of which 8 studies remained for extraction and analysis. More detail is provided in the PRISMA flow diagram in Figure 1, including reasons for exclusion at the full-text phase. The key study characteristics for all included papers are presented in Table 2. A summary of key psychometric validation results can be found in Multimedia Appendix 3 [15,47,48]. All studies included psychometric validity testing as part of their analysis except the Virtual Supermarket Test (VST) [37], where validity testing had been completed previously [15,47,48]. All games were designed and developed by the researchers involved in the included papers, with the exception of the Smart Aging Serious Game (SASG) [38], which was developed by Zucchella et al [49]. None of the included papers involved testing commercially available games; however, 2 of the games [39,40] were experimentally controlled versions of existing games. In addition, 4 of the 8 game studies were noted by the authors as being “virtual reality” interventions, which attempted to replicate real-life activities to assess cognition. However, all these games had low immersion and adhered to our definition of a digital GBCA (refer to the Discussion of important terms subsection for further justification).

Figure 1. PRISMA flow diagram.

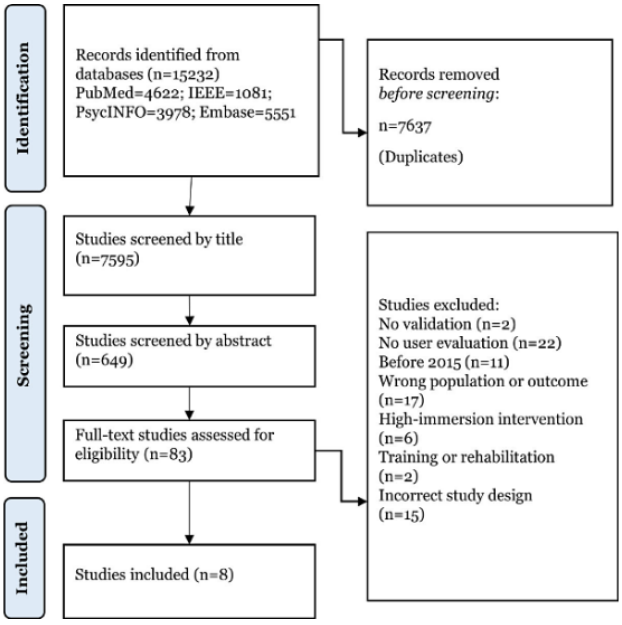


Table 2. Summary of key study characteristics.

Studies	Game	Sample, n	Age, mean	Female (%)	Study aim
Cabinio et al [38], 2020	SASG[a]	Total: 139; controls: 107; MCI[b]: 32	Controls: 76.47; MCI: 76.75	Controls: 49.5; MCI: 46.8	Tested the usability of SASG in a cohort of participants with MCI and determined the validity of the SASG in discriminating between a preclinical population with MCI and healthy controls.
Chesham et al [40], 2019	SMT[c]	Total: 52; young (18-35 y): 28; older (65-85 y): 13; oldest (≥85 y): 11	Young: 21.68; older: 70.54; oldest: 89.27	Young: 71.4; older: 53.9; oldest: 81.8	Examined the initial validity and usability of an experimentally controlled version of a popular TMM3 puzzle video game—the SMT.
Manera et al [41], 2015	Kitchen and Cooking	Total: 21; MCI: 9; AD[d]: 12	MCI: 75.8; AD: 80.3	MCI: 77.8; AD: 86.7	Conducted a feasibility study of patients with MCI and AD and related disorders using the game "Kitchen and Cooking."
Nef et al [39], 2020	Numberlink puzzle task	Total: 55; young (18-31 y): 18; older (64-79 y): 14; oldest (86-98 y): 14; PD[e]: 4; HD[f]: 5	Young:21.83; older: 71.4; oldest: 89.4; PD: 67.5; HD: 49.4	Young: 66.7; older: 57.1; oldest: 85.7; PD: 50; HD: 50	Evaluated the feasibility and preliminary validity of a maze-like NL puzzle video game as a tool to assess cognitive and motor differences in adults and patients with neurodegenerative disorders.
Valladares-Rodriguez et al [16], 2017	Episodix	Total: 16; controls: 8; MCI: 3; AD: 5	Controls: 68.3; MCI: 75.8; AD: 75	Total: 75	Investigated if a video game could be designed and developed to assess episodic memory and predict early cognitive impairments in an ecologic setting with low immersion.
Vallejo et al [42], 2017	Virtual games	Total: 38; controls: 20; AD: 18	Controls: 74.6; AD: 77.8	Controls: 40; AD: 50	Evaluated the usability and the screening potential of several low-immersion virtual game tasks for patients with AD.
Wang et al [43], 2022	GBCA[g]	Total: 124; controls: 57; NCDs[h]: 67	Controls: 72.31; NCDs: 74.78	Controls: 69; NCDs: 63	Developed a game-based tool and evaluated its validity as an early screening for patients with cognitive impairment.
Zygouris et al [37], 2022	VST[i]	Total: 57; SCD[j]: 24; MCI: 33	SCD: 65.58; MCI: 68.39	SCD: 79; MCI:70	Assessed the VST's usability in a sample of older adults with MCI and SCD. The paper analyzed how usability is affected by age, education, diagnosis, in-game performance, and familiarity with touch devices.

[a]SASG: Smart Aging Serious Game.

[b]MCI: mild cognitive impairment.

[c]SMT: search and match task.

[d]AD: Alzheimer's disease.

[e]PD: Parkinson disease.

[f]HD: Huntington disease.

[g]GBCA: game-based cognitive assessment.

[h]NCD: neurocognitive disorder.

[i]VST: Virtual Supermarket Test.

[j]SCD: subjective cog decline.

User Evaluation Findings

Overview

Preliminary analysis of the included papers identified 2 distinct areas of focus: usability and user experience. As such, our proceeding synthesis is demarcated based on usability and user experience (applying the simple definitions of these constructs provided in the Introduction section). Reporting findings from these 2 overlapping yet distinctive areas provides a holistic perspective of user evaluation by incorporating both system-level usability results as well as more individualized user experiences of serious gameplay. A detailed overview of relevant study findings and key statistical results is presented in Multimedia Appendix 4 [50-56].

GBCA Usability

A total of 5 studies [37-40,42] included system-level usability results. All these studies involved participant self-report measures but varied in the level of detail participants were asked to report. Overall, the average usability, from a systems perspective, of the games presented ranged from good to excellent. Familiarity with computers or touch screens did not appear to have a significant effect on system usability for computers or touchscreen devices, which is a positive finding. However, there were distinct differences in system usability across age groups and cognitive functioning levels in some

studies. An interesting relationship was also found between usability and game performance in the VST study [37]. These are discussed further in the subsequent sections.

In the SASG study [38], participants were simply asked to respond to some questions about their familiarity with computers and touch screens. Results showed that familiarity with computer systems did not influence the SASG game performance score. Although this is not in itself an assessment of usability, it revealed an important usability finding: a lack of digital experience did not impact overall cognitive performance scores in a sample of older game users with low digital literacy and experience. A potential reason for this result may be the cautious and simplistic design of the SASG interface, which supported usability for the target cohort. The game was built specifically for older and nonexpert users, and according to the authors, it did not require skilled digital abilities. Using a low-immersion, first-person perspective, the game required users to complete gamified versions of everyday tasks, such as watering flowers with the radio on and dialing a phone number. Furthermore, the authors suggested a key factor in the usability of their serious game intervention was the decision to use a touch screen instead of a mouse.

The remaining 4 studies used the system usability scale (SUS) to measure and analyze usability [57]. The SUS is a brief questionnaire (10 questions) using a 5-point Likert scale with a possible aggregate score ranging from 0 to 100. The SUS includes questions such as "I thought the system was easy to use" and "I needed to learn a lot of things before I could get going with this system." According to previous studies [58,59], the average SUS score (50th percentile) is 68, a good score is between 71 and 84, and an excellent usability score is ≥85.

In the VST study [37], the average SUS score was 83.11 (SD14.6). No statistically significant differences were found between the participant groups, people with subjective cognitive impairment and mild cognitive impairment (MCI), regarding SUS scores. However, there was a significant correlation (r=−0.496; P<.001) found between the SUS score and the average time needed to complete the VST test trials, independent of the participant group. This suggested a relationship between game completion and usability: the lower the participants rated the game's usability, the longer they took to complete it. Although this was an interesting finding, the direction of causality is unknown. As the authors noted [37], "It is unclear whether participants who took longer to complete the VST test trials provided a lower usability score as they were frustrated by the long time it took them to complete the exercise or if there are underlying usability issues that affected their performance thus resulting in a longer time to complete the VST test trials." It is also worth noting that in the VST study, participants were asked to complete the VST exercise on their own. The examiner prepared the tablet and launched the VST game but did not assist the participants. In fact, researchers intentionally provided little support to participants once gameplay started. The reason for this was to assess the usability of the system in an environment where a user may need to self-administer the game without support.

In the Numberlink (NL) puzzle task [39], the SUS average was 93.38 (SD 5.72) for young adults, 93.33 (SD 6.25) for older adults, 83.39 (SD 11.79) for the oldest adults, 81.88 (SD 9.66) for people with PD, and 83.12 (SD 10.68) for those with HD. These scores reflected a statistically significant difference in usability ratings between groups (P=.02). There were similar findings in the search and match task (SMT) study, albeit with lower reported usability overall [40]. The average SUS score was 88.67 (SD 7.28) for younger adults, 79.09 (SD 15.50) for older adults, and 68.25 (SD 18.78) for the oldest adults. This indicated a significant effect of age group (P=.01), with usability ratings decreasing with increasing participant age. Finally, the virtual game study [42] reported a mean SUS score of 83.5 (SD 11.16) for the control group and 83.75 (SD 9.82) for people with AD. As such, similar to the VST study and contrary to the NL puzzle task and SMT studies, there was no significant group difference found in usability. There was also no significant difference in SUS scores found between participants with and without touch device familiarity.

GBCA User Experience

Five studies [16,39-41,43] examined user experiences of gameplay, specifically focusing on participant engagement with game elements, perceived motivation, and enjoyment. Overall, all the games presented were well received by users regarding motivation, enjoyability, and satisfaction. However, there were some important differences across the study findings. For instance, some studies identified a relationship between being cognitively healthy and improved user experience [40,43], whereas an opposite relationship was found in another study [41]. Furthermore, 2 studies identified the role of a health professional as essential to the success of their respective game interventions [16,41], which contrasted with the intent in studies to test implementing games in a self-administering environment [37,41].

In the Episodix study [16], a user experience questionnaire based on game usefulness and user motivation was completed by participants. The questionnaire was completed twice, once before and once after gameplay. Before playing the game, most participants reported low interest, motivation, and perceived usefulness of serious games; however, their experiences playing Episodix changed these perceptions. For instance, the average participant's motivation to play video games increased by 40% after gameplay. Participants also indicated that the Episodix game seemed more engaging than the California Verbal Learning Test (CVLT) pen-and-paper test, which the game was based on. One potential reason for the game seeming more engaging than the CVLT may have been the comprehensive user-centered design process used in developing the game. A combination of validity testing, expert input, and user focus groups underpinned the development process. Through the design process, participants from both focus groups and the pilot experiment indicated a preference for a touch interface rather than a traditional keyboard and mouse setup. This finding was similar to that of the SASG game study. Technical developments were made to facilitate these preferences. The Episodix study [16] also reported important qualitative findings from user focus groups. The game was perceived as very useful by users, but they preferred it to be administered by a health

professional rather than self-administered, regardless of whether they were confident interacting with the game on their own. Participants agreed that if both instruments (the CVLT and the game) were demonstrated to perform the same cognitive assessment, they would prefer to play the game. However, participants preferred the digital intervention to be referred to as a test because they perceived the term *video game* to have negative and nonserious connotations.

Both the NL puzzle task and SMT studies [39,40] assessed user experience through the Perception of Game Training Questionnaire [50]. Participants rated the extent to which they found playing the games enjoyable, challenging, frustrating, and motivating. Both games were developed based on experimentally controlled versions of commercially available puzzle games. In addition, both games offered different difficulty levels to create an appropriate level of challenge for a broad range of users. In the NL puzzle task study, ratings of enjoyment (P=.43), challenge (P=.07), frustration (P=.06), and motivation (P=.67) for the game did not differ significantly between groups. The users generally rated the game experience positively in both the healthy aging group and the group with cognitive impairment. However, there was a slight, albeit statistically nonsignificant, increase in rating the game session as challenging and frustrating among older participants and/or those with impairment, particularly participants with PD and HD. According to the authors, further adapting the difficulty levels to a person's performance might partially reverse this trend. In the SMT study, there were significant differences between the 3 age groups regarding ratings of challenge (P<.001). Overall, young adults perceived the SMT as significantly less challenging than older adults (P=.02). There were also significant age group differences in average difficulty ratings for both the short version played by the young, older, and oldest adults (P<.001) and the long version (P<.001) played by the young and older adults. Despite this, there were no significant group differences regarding enjoyment, frustration, and motivation while playing the SMT task. One possible reason why challenge and difficulty levels varied across groups while enjoyment and motivation levels remained constant is that the SMT provided a total of 71 possible difficulty levels. This means that the game could be responsive to a user's gameplay and adjust the level of perceived challenge accordingly.

In the GBCA study [43], the user experience questionnaire included questions on game ease-of-use (not technically user experience), whether users felt the game stories were interesting and familiar, users' reactions to the interface design (typesetting, color, instruction, etc), and questions aimed at comparing the GBCA intervention with other cognitive assessment tools. The responses of the control group suggested that they were very satisfied with the game. By contrast, the average responses of participants with neurocognitive disorders were significantly lower than those of controls. Interestingly, lower scores given by the participants on the user experience questionnaire corresponded to a higher clinical dementia rating, older age, and a lower educational level. Higher levels of cognitive impairment appeared to adversely affect the participants' ability to use the tablet to complete the GBCA intervention, as indicated by the poor ratings given to questions regarding the design of the GBCA intervention. This finding raises questions about the functional usability of the game as well as the more subjective user experience. Furthermore, older age had a negative influence on the participants' satisfaction regarding the pictures, text, and story used in the game. This was somewhat surprising, given that the user interface of the GBCA intervention was intentionally designed to be acceptable for older users and included large pictures and buttons as well as spoken instructions so users could operate the app easily. The game could also be played by tapping buttons, drawing by dragging one's fingers over the touch screen, and speaking into a microphone (using low-cost speech recognition technology).

The Kitchen and Cooking study [41] also captured user experiences. The participants completed a questionnaire about the game experience at 2 points during the study. Specifically, satisfaction, interest, motivation (intrinsic and extrinsic), emotional experience of gameplay, and fatigue levels were assessed. The overall results of the self-report questionnaire showed that the participants were highly satisfied, interested, and motivated by the game experience. Intrinsic motivation was significantly higher than extrinsic motivation. Average user experience scores did not change between the 2 sessions, thus confirming the overall positive evaluation of the game after repeated gameplay. However, it is worth noting that scores did not trend upward after the second session regarding motivation, as observed in the Episodix study [16]. Of particular interest in this study is that participants with AD reported being significantly more satisfied with the game than participants with MCI (P=.04). This contrasts with the findings from the NL puzzle task, SMT, and GBCA studies, where the inverse trend appeared. Finally, the authors suggested that a critical factor in the success of the intervention was the presence of a clinician. The game was designed as a tool to help patients train outside clinical consultations. However, the periodic supervision of the clinician helped to explain the functioning of the game to the patients and their families, keep track of the evolution of the performance, adapt the intervention to the patients' changing needs, and maintain user motivation.

Discussion

Overview

The aim of this review was to investigate studies that evaluated the experiences of older people using GBCAs. Although the combined findings are from only 8 included papers, they establish promising preliminary user evaluation results for the use of GBCAs, albeit with some interesting implications for future design and development practices. Our review provides a much-needed synthesis of user evaluation findings from GBCAs, which have undergone psychometric validity testing. As such, we offer new insights regarding the relationship between the validity, usability, and user experience of GBCAs. This is particularly important with respect to the experiences of our target study cohort, older people (aged ≥50 years), whose input is greatly needed yet seldom included in GBCA design and development.

The first area of focus in this review was system usability. Usability scores were mostly positive across the studies

included. However, in some of the games presented, there were trends toward lower usability scores for participants with cognitive impairment. This was observed even when games were explicitly designed for older people and/or those with impairment, and tasks were simplistic and representative of basic daily activities, such as in the VST. In the VST, a game that has been validated in multiple contexts [15,47,48], no significant differences were found between the groups with subjective cognitive impairment and MCI in usability scores; however, participants who took longer to complete the test trials had lower self-reported usability. Usability challenges were further identified in the NL puzzle task and the SMT. Although both games had good usability scores on average, they both identified significant differences in usability ratings not only across cognitive health but also age range. On average, in both games, the older or more impaired a person was, the less usable they found the respective game system.

These collective findings on usability have some interesting implications. Traditionally, it has been assumed that system usability can be increased by improving the technical interactions a person has with a system, the intuitiveness of that system, and by reducing any bugs or errors that generate needless complexity within that system. For example, Joddrell and Astell [60] significantly improved the usability of solitaire for people with dementia (mean age 84.17, SD 8.35; range 66-102) by adding accessibility features such as additional card control options, clearer gameplay layout, and audio-visual feedback cues. However, this review highlights that while these technical and functional system improvements may be important, they are not necessarily sufficient to produce usable interventions for this target group of users. Individual variations in age, cognitive function, and gameplay performance may all impact functional usability, even when the system has been technically adapted to be suitable for older users and/or those with impairment.

The second focus area, regarding subjective user experience, identified the importance of perceived challenge in mediating gameplay experience across groups. In the GBCA study, the healthy control group was more satisfied with their experience of using the GBCA intervention than the group of participants with cognitive impairment, whereas the opposite trend appeared in the Kitchen and Cooking game. In the Kitchen and Cooking game, a simplistic interface based on a basic cooking task, game satisfaction was higher among the participants with AD than among participants with MCI. In the SMT task, there was a statistically significant difference in how challenging older and younger people found the game. However, levels of enjoyment and motivation remained the same. These contrasting results may be related to perceived gameplay challenges. For instance, the NL puzzle task, the SMT, and the GBCA were, on average, more complex than the Kitchen and Cooking game, with the NL puzzle task and SMT offering different levels of complexity for different users.

According to the broader literature on game design preferences, people want to be challenged to a level that will satisfy their need for competence [25]. Applying self-determination theory, it has been shown that a player will experience low competence if the challenges are too great (eg, game controls are overly complex or enemies are too numerous), and this will have a negative impact on a person's ability to enjoy the game [61]. If a game is too difficult, it can become frustrating and discouraging, reducing a user's motivation, engagement, and enjoyment. The optimal challenge (ie, the sweet spot) for any player is mediated by that individual's threshold for difficulty [62,63]. This may explain why basic gamified versions of simple tasks, such as in the Kitchen and Cooking game, were less motivating or enjoyable for younger participants and those without cognitive impairment. It may also explain why more "from scratch" serious games that offered more difficulty levels, such as the NL puzzle task and SMT, had constantly high enjoyment and motivation scores across the younger, older, and oldest groups of users.

Another noteworthy finding in this review relates to the role of a health professional during game-based assessment. Two studies identified the role of a health professional as essential to the success of their respective game interventions. In the Episodix game, although users demonstrated improved appreciation for the value of GBCAs after gameplay, this increased appreciation appeared to stem from the value of the game as a health tool. The presence of a clinician seemed to underpin this, as the game was perceived as very useful by users, but they preferred it to be administered by a health professional rather than self-administered. Likewise, in the Kitchen and Cooking study, the authors suggested a critical success factor for the intervention was the presence of a clinician. These findings raise questions about self-administration, which is seen as one of the most exciting (and valuable) applications of serious game technology [15]. Self-administration of GBCAs has been recently identified as a way to overcome some of the issues with formal assessment [15], such as significant test anxiety as well as self-stigma regarding low literacy or education levels when performing cognitive tests [9,10]. Self-administration may also empower people to take control of their own health earlier and more proactively than standard approaches. However, the findings of this review suggest that this autonomy and control may come with some unintended consequences, such as a lack of confidence in undertaking assessments without support, as well as a lower perceived value or legitimacy of the assessment. This is important to recognize as a potential downside for self-administrated GBCAs.

The Objectives of GBCAs and Subsequent Tensions With Usability and User Experience

One approach often taken when designing and developing cognitive games for older people and/or individuals with cognitive impairment is to make the game easy. This appears to have been the design strategy in games such as SASG, VST, and Kitchen and Cooking, where the games were based on simple daily living tasks, and the design and development of the games were tailored to people with cognitive decline. These games were quite successful at engaging older people with significant cognitive impairment. However, given that the objective of a GBCA is to assess maximal cognitive capacity, adapting the floor and ceiling to make games easier may not always be desirable. This is because the reason for adding complexity to a GBCA, besides increasing user engagement, is to improve the sensitivity of a game to detect cognitive

functioning more accurately. In other words, a game that is too simple or too domain specific may not be relevant for healthier older adults or those with milder or more complex forms of cognitive impairment, even if it is more appropriate for people with more significant cognitive decline. For instance, the basic requirement of shopping for groceries in the VST may be suitable for people with moderate neurocognitive disorders; however, it may not detect subtle impairments or small cognitive changes over time in healthier participants. More simplistic games are also unlikely to be very enjoyable or challenging (ie, satisfying a need for competence) for many people without neurocognitive disorders (as was observed in some of the studies included). This reduces their viability as a scalable screening tool. By adding more challenge and complexity, it becomes feasible to engage a broader user group and assess multiple cognitive domains. This, of course, comes at the risk of marginalizing the most vulnerable and arguably important cohort: those with significant impairments. It is also important to note that people with more significant neurocognitive disorders are unlikely to be the intended target of most GBCAs, or even traditional cognitive screening. This is a tension that has been seldom discussed or investigated in the growing body of literature on GBCAs.

Overall, the goals of usability and user experience—foundational to human-centered design approaches—and the intention of GBCAs and cognitive testing, in general, reveal a potentially problematic misalignment that warrants further discussion. For instance, according to ISO, the basic requirement of usability is that a product can be used (by target users) to achieve specific goals with effectiveness, efficiency (ie, in minimum time and with minimum errors), and satisfaction in a specified context [27]. However, is it reasonable, desirable, or even possible to meet this standard when developing a GBCA? If the primary goal of a cognitive test is the detection of cognitive deficits, the implication is that people with cognitive impairment will likely (1) fail to achieve certain tasks, (2) take longer to complete objectives, and (3) make more errors than cognitively healthy people.

In fact, reliable detection of these "failures" is essential to ensure a valid and clinically meaningful test of cognitive performance. Thus, intentionally producing user failure states is a central design tenet when developing GBCAs, and this, as we have identified through this review, may cause reductions in usability and subjective user experiences.

However, while this may risk departing somewhat from the principles of user-centered design, it is important to recognize that the "serious" objective of a GBCA is not necessarily to maximize unconstrained usability and user experience; rather, it is to produce a cognitive assessment that is more motivating and engaging than a traditional task. As traditional tasks are sometimes reported to be tedious, boring, anxiety provoking, and stigmatizing, the objective of a GBCA is to improve the subjective user experience relative to traditional approaches while maintaining functional usability. This implies that a GBCA may not necessarily need to be maximally fun but simply "fun enough" while being motivating, engaging, and technically usable. Furthermore, what "usable" means in this context is potentially hard to reconcile with the ISO aim of effectiveness, efficiency, and satisfaction. For instance, if an intervention is too easy to use, can it adequately produce the failure states that are the primary objective of a GBCA?

These open questions become even more interesting when we consider that the intended application of GBCAs, at least in the near future, is likely to be a relatively short (eg, 10-20 min) user experience (ie, replacing a cognitive screen), which is only "played" by users sporadically. The implication is that traditional user-centered game design that focuses on intrinsically motivating features, such as narrative and social connectedness, may not be feasible or desirable. This, in many ways, is misaligned with the overarching goals of human-centered design, yet it is arguably more consistent with the preferences of older people taking a cognitive test. In fact, this was what we reported in a recent qualitative study investigating the GBCA preferences of older people in prison [25]. We found that GBCAs with numerous immersive game features risked being perceived as too childish for the serious context of a cognitive test. This was also likely the case in the Episodix study sample, where participants preferred the digital intervention to be referred to as a test because they perceived the term game to be nonserious [16]. However, we argue that these tensions do not reduce the value of incorporating usability and user experience principles into the development of GBCAs; rather, they simply reframe what the feasible intention of a GBCA ought to be in the context of user experience and usability. Perhaps "good enough" usability and user experience are good enough?

Our conclusion is that there is an inherent tension between usability, user experience, and GBCAs that (1) in the meantime should be carefully considered when designing and developing GBCAs, especially for older people; and (2) requires additional research to further explore and potentially resolve. We offer some key lessons in the Implications for Practice subsection. Regarding future research, we suggest an urgent need for further mixed methods research that incorporates the qualitative preferences of older people using GBCAs with rigorous validity testing. In addition, there is a need to explore what "good enough" usability and user experience may mean in this context and how these concepts can be tested and incorporated into game design, development, and improvement.

Limitations

This review has some limitations. Many of the participant samples were small, and the studies involved slightly different target user groups. Most studies focused on user evaluation as a secondary outcome, with psychometric validity testing being the primary aim in most papers. All the games included in this review were distinct from one another regarding design and development processes and, as a result, the end products. Combined with the disparities in sampled participants, this heterogeneity makes it inappropriate to make specific recommendations regarding which precise techniques work for whom in what contexts. This is further complicated by the fact that most of the games did not seem to be publicly available, creating uncertainty about how well developed and sophisticated the games actually are in practice. This lack of access to the games, beyond what was described and visualized in the primary research papers, also inhibited a more comprehensive review

of specific design choices based on a more technical analysis of mechanics and processes. However, despite these limitations, we were still able to interpret the findings in the context of the broader literature on serious games and provide general lessons learned for designing and developing GBCAs.

In addition, there were also some limitations related to the methods applied for user evaluation. Most user evaluation data were collected via self-report and are thus exposed to the well-established biases associated with this collection method. Furthermore, although the SUS is well validated, it is short and generic. Thus, it does not present a comprehensive assessment of game usability, and its validity for people with neurocognitive disorders has not been established. Other reported usability and user experience findings came from disparate tools, some of which were customized from other scales or created based on theory. This created additional barriers when attempting to synthesize the findings quantitatively and prevented a meta-analysis. Some of these limitations were evident in the quality appraisal, in which a case-control checklist was used to assess the risk of bias. This basic checklist was deemed appropriate because most of the study papers included games that were still in the pilot stage of development and, thus, not ready for more sophisticated diagnostic appraisal. Regardless, it is hoped that a more sophisticated and suitable appraisal tool can be used in future research to adequately assess the quality of papers investigating the psychometric validity and user experiences of GBCAs.

Implications for Practice

Overview

This systematic review aimed to synthesize the user evaluation data from psychometrically validated GBCAs played by older people. As part of this process, we have identified key lessons for developing potentially transformative assessments using gamified technology. These preliminary key lessons have been developed based on our systematic review of the included studies, our own primary research experiences (eg, Mantell et al [25]), and a broader synthesis of relevant theory and literature (eg, self-determination theory). Given that this is still a relatively new area of scientific inquiry, we hope that these lessons can serve as guidance for clinicians, researchers, and game developers who find themselves navigating the complexity of creating GBCAs that are both clinically effective and offer "good enough" user experiences and usability. More broadly, we hope this review can support and inform more evidence-based and rigorous design, development, and evaluation of GBCAs for older users in the future.

Validity Is Very Necessary but Not Sufficient

A game may function as an assessment tool in theory, but if it is not usable or engaging, it is unlikely to work in practice. Input from both users and game designers, at the outset and throughout the design process, is the best way to ensure a game is suitable. This is usually best done alongside an expert reference group working to ensure the game has psychometric validity [13]. The more the target users, developers, and cognitive experts collaborate, the more likely the game will succeed. Furthermore, researchers should always plan to evaluate game usability, as well as user experiences and preferences, to formally investigate the success of their game (and identify any scope for improvement).

Clarify the Intended User

There is a complex relationship between cognitive decline, usability, and user experience as people age. Many older people with cognitive impairment will respond better to simple gamification of realistic tasks. More complex serious games risk becoming less appealing as people become older or more impaired, and this will impact usability and user experience. However, the more rudimentary the game-based task, the less likely it is to engage younger adults or those with better cognitive function. To proactively address this tension, developers should begin with a critical question: who is this intervention for, and what exactly is the game intended to assess? For instance, "good enough" usability and user experience are probably good enough for a short and sporadic cognitive screener targeting older adults at risk of impairment. However, this may not be the case for a more expansive game trying to also engage younger cohorts and track more subtle cognitive changes over time.

Create Games That Align With the Unique Preferences and Needs of Older People

Older people appear to prefer games that avoid childlike features, are intellectually focused (eg, puzzles), and/or are based on the replication of real-life tasks or challenges. Avoiding fantasy, violent content, and multiplayer interaction is recommended. Large digital touchscreen devices, such as tablets, also seem to be more usable and intuitive for older people than smaller devices, mouses, or virtual reality [16,38,64]. However, tablets can also create ergonomic difficulties for older people who may be frail or have specific neurological conditions (eg, difficulty holding the device steadily, dyspraxia, or visuospatial impairment), and these factors need to be accounted for [18].

Providing Each User With Their Optimal Challenge Is Key

Tailoring the level of difficulty (known as difficulty balancing) to an individual's capacity should be attempted wherever possible in the design of GBCAs. It has been shown that this personalization does not need to be highly sophisticated to be useful in increasing user experience [37,39]. Introducing various levels or randomly generating different design elements or challenges per round can also make it possible to detect more subtle impairments and reduce learning effects [37]. In addition, it can also alleviate some of the tension between providing games that are suitable for older people or those with more impairment and maintaining sensitivity for more subtle cognitive decline, as more levels mean a lower floor and higher ceiling.

Acknowledgments

This work was funded through an Australian National Health and Medical Research Council Ideas Grant (grant GA143456; ASCAPE: Audio App-delivered Screening for Cognition and Age-related Health in Older Prisoners).

Conflicts of Interest

MK is the chief executive officer of Arludo.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 22 KB - aging_v8i1e65252_app1.docx]

Multimedia Appendix 2
Risk of bias.
[DOCX File , 24 KB - aging_v8i1e65252_app2.docx]

Multimedia Appendix 3
Validity summary.
[DOCX File , 26 KB - aging_v8i1e65252_app3.docx]

Multimedia Appendix 4
User evaluation summary.
[DOCX File , 25 KB - aging_v8i1e65252_app4.docx]

References

1. Friehs MA, Dechant M, Vedress S, Frings C, Mandryk RL. Effective gamification of the stop-signal task: two controlled laboratory experiments. JMIR Serious Games 2020 Sep 08;8(3):e17810 [FREE Full text] [doi: 10.2196/17810] [Medline: 32897233]
2. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 2016 Jul 15;4(2):e11 [FREE Full text] [doi: 10.2196/games.5888] [Medline: 27421244]
3. Wiley K, Robinson R, Mandryk RL. The making and evaluation of digital games used for the assessment of attention: systematic review. JMIR Serious Games 2021 Aug 09;9(3):e26449 [FREE Full text] [doi: 10.2196/26449] [Medline: 34383674]
4. About dementia. Dementia-Australia. URL: https://www.dementia.org.au/information/about-dementia [accessed 2025-05-29]
5. Manning V. Managing cognitive impairment in AOD treatment guidelines. Turning Point. URL: https://www.turningpoint.org.au/treatment/clinicians/Managing-Cognitive-Impairment-in-AOD-Treatment-Guidelines [accessed 2025-05-29]
6. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
7. Pedersen MK, Díaz CM, Wang QJ, Alba-Marrugo MA, Amidi A, Basaiawmoit RV, et al. Measuring cognitive abilities in the wild: validating a population-scale game-based cognitive assessment. Cogn Sci 2023 Jun;47(6):e13308. [doi: 10.1111/cogs.13308] [Medline: 37354036]
8. National Collaborating Centre for Mental Health. Dementia: A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care. London, UK: British Psychological Society; 2007.
9. Dorenkamp MA, Vik P. Assessment-related anxiety among older adults: development of a measure. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Sep 27;28(5):781-795. [doi: 10.1080/13825585.2020.1826398] [Medline: 32985347]
10. Dorenkamp MA, Vik PW. Neuropsychological assessment anxiety: a systematic review. Pract Innov 2018 Sep;3(3):192-211. [doi: 10.1037/pri0000073]
11. Byrd DA, Rivera Mindt MM, Clark US, Clarke Y, Thames AD, Gammada EZ, et al. Creating an antiracist psychology by addressing professional complicity in psychological assessment. Psychol Assess 2021 Mar;33(3):279-285. [doi: 10.1037/pas0000993] [Medline: 33779204]
12. Tong T, Chignell M, Tierney MC, Lee J. A serious game for clinical assessment of cognitive status: validation study. JMIR Serious Games 2016 May 27;4(1):e7 [FREE Full text] [doi: 10.2196/games.5006] [Medline: 27234145]
13. Khaleghi A, Aghaei Z, Mahdavi MA. A gamification framework for cognitive assessment and cognitive training: qualitative study. JMIR Serious Games 2021 May 18;9(2):e21900 [FREE Full text] [doi: 10.2196/21900] [Medline: 33819164]

14. Leduc-McNiven K, White B, Zheng H, D McLeod R, R Friesen M. Serious games to assess mild cognitive impairment: 'the game is the assessment'. Res Rev Insights 2018;2(1):52. [doi: 10.15761/rri.1000128]
15. Zygouris S, Iliadou P, Lazarou E, Giakoumis D, Votis K, Alexiadis A, et al. Detection of mild cognitive impairment in an at-risk group of older adults: can a novel self-administered serious game-based screening test improve diagnostic accuracy? J Alzheimers Dis 2020;78(1):405-412 [FREE Full text] [doi: 10.3233/JAD-200880] [Medline: 32986676]
16. Valladares-Rodriguez S, Perez-Rodriguez R, Facal D, Fernandez-Iglesias MJ, Anido-Rifon L, Mouriño-Garcia M. Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. PeerJ 2017;5:e3508 [FREE Full text] [doi: 10.7717/peerj.3508] [Medline: 28674661]
17. Birk MV, Mandryk RL, Bowey J, Buttlar B. The effects of adding premise and backstory to psychological tasks. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '15; April 18-23, 2015; Seoul, Republic of Korea p. 2015 URL: http://gamification-research.org/wp-content/uploads/2014/11/GAMICHI15_birk_et_al.pdf
18. Tong T, Chignell M, Lam P, Tierney MC, Lee J. Designing serious games for cognitive assessment of the elderly. Proc Int Sympt Hum Factors Ergon Health Care 2014 Jul 22;3(1):28-35. [doi: 10.1177/2327857914031004]
19. Ferreira-Brito F, Fialho M, Virgolino A, Neves I, Miranda AC, Sousa-Santos N, et al. Game-based interventions for neuropsychological assessment, training and rehabilitation: which game-elements to use? A systematic review. J Biomed Inform 2019 Oct;98:103287 [FREE Full text] [doi: 10.1016/j.jbi.2019.103287] [Medline: 31518700]
20. Mantell R, Hwang YI, Radford K, Perkovic S, Cullen P, Withall A. Accelerated aging in people experiencing homelessness: a rapid review of frailty prevalence and determinants. Front Public Health 2023 Mar 16;11:1086215 [FREE Full text] [doi: 10.3389/fpubh.2023.1086215] [Medline: 37006541]
21. Withall A, Mantell R, Hwang YI. Background paper: issues facing older people who are leaving prison. Australian Association of Gerontology. 2022. URL: https://aag.asn.au/libraryviewer?ResourceID=74 [accessed 2025-05-29]
22. Astell AJ, Joddrell P, Groenewoud H, de Lange J, Goumans M, Cordia A, et al. Does familiarity affect the enjoyment of touchscreen games for people with dementia? Int J Med Inform 2016 Jul;91:e1-e8. [doi: 10.1016/j.ijmedinf.2016.02.001] [Medline: 26897552]
23. Chesham A, Wyss P, Müri RM, Mosimann UP, Nef T. What older people like to play: genre preferences and acceptance of casual games. JMIR Serious Games 2017 Apr 18;5(2):e8 [FREE Full text] [doi: 10.2196/games.7025] [Medline: 28420601]
24. Blocker KA, Wright TJ, Boot WR. Gaming preferences of aging generations. Gerontechnology 2014;12(3):174-184 [FREE Full text] [doi: 10.4017/gt.2014.12.3.008.00] [Medline: 29033699]
25. Mantell R, Withall A, Radford K, Kasumovic M, Monds L, Hwang YI. Design preferences for a serious game-based cognitive assessment of older adults in prison: thematic analysis. JMIR Serious Games 2023 Apr 17;11:e45467 [FREE Full text] [doi: 10.2196/45467] [Medline: 37067850]
26. de Paula DF, Menezes BH, Araújo CC. Building a quality mobile application: a user-centered study focusing on design thinking, user experience and usability. In: Proceedings of the 3rd International Conference on Design, User Experience, and Usability: User Experience Design for Diverse Interaction Platforms and Environments. 2014 Presented at: DUXU '14; June 22-27, 2014; Crete, Greece p. 313-322 URL: https://link.springer.com/chapter/10.1007/978-3-319-07626-3_29 [doi: 10.1007/978-3-319-07626-3_29]
27. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. 2024. URL: https://www.iso.org/standard/77520.html [accessed 2025-05-29]
28. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
29. Gaver B, Martin H. Alternatives: exploring information appliances through conceptual design proposals. In: Proceedings of the 2000 SIGCHI conference on Human Factors in Computing Systems. 2000 Presented at: CHI '00; April 1-6, 2000; The Hague, The Netherlands p. 209-216 URL: https://dl.acm.org/doi/10.1145/332040.332433 [doi: 10.1145/332040.332433]
30. Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the 1994 ACM Conference on Human Factors in Computer Systems. 1994 Presented at: CHI '94; April 24-28, 1994; Boston, MA p. 152-158 URL: https://static.aminer.org/pdf/PDF/000/089/679/enhancing_the_explanatory_power_of_usability_heuristics.pdf
31. Kletzel SL, Sood P, Negm A, Heyn PC, Krishnan S, Machtinger J, et al. Effectiveness of brain gaming in older adults with cognitive impairments: a systematic review and meta-analysis. J Am Med Dir Assoc 2021 Nov;22(11):2281-8.e5 [FREE Full text] [doi: 10.1016/j.jamda.2021.05.022] [Medline: 34146517]
32. Ning H, Li R, Ye X, Zhang Y, Liu L. A review on serious games for dementia care in ageing societies. IEEE J Transl Eng Health Med 2020;8:1-11. [doi: 10.1109/jtehm.2020.2998055]
33. Sabermahani F, Almasi-Dooghaee M, Sheikhtaheri A. Alzheimer's disease and mild cognitive impairment serious games: a systematic analysis in smartphone application markets. Stud Health Technol Inform 2022 May 16;293:179-186. [doi: 10.3233/SHTI220366] [Medline: 35592979]
34. Valladares-Rodríguez S, Pérez-Rodríguez R, Anido-Rifón L, Fernández-Iglesias M. Trends on the application of serious games to neuropsychological evaluation: a scoping review. J Biomed Inform 2016 Dec;64:296-319 [FREE Full text] [doi: 10.1016/j.jbi.2016.10.019] [Medline: 27815228]

35. Mezrar S, Bendella F. A systematic review of serious games relating to cognitive impairment and dementia. J Digit Inf Manag 2022;20:1. [doi: 10.6025/jdim/2022/20/1/01-09]
36. Sleith C. Methodology checklist 4: case control studies healthcare improvement Scotland. Scottish Intercollegiate Guidelines Network. 2012. URL: https://www.sign.ac.uk/using-our-guidelines/methodology/checklists/ [accessed 2025-05-29]
37. Zygouris S, Segkouli S, Triantafyllidis A, Giakoumis D, Tsolaki M, Votis K, et al. Usability of the virtual supermarket test for older adults with and without cognitive impairment. J Alzheimers Dis Rep 2022 May 02;6(1):229-234. [doi: 10.3233/adr-210064]
38. Cabinio M, Rossetto F, Isernia S, Saibene FL, Di Cesare M, Borgnis F, et al. The use of a virtual reality platform for the assessment of the memory decline and the hippocampal neural injury in subjects with mild cognitive impairment: the validity of smart aging serious game (SASG). J Clin Med 2020 May 06;9(5):1355 [FREE Full text] [doi: 10.3390/jcm9051355] [Medline: 32384591]
39. Nef T, Chesham A, Schütz N, Botros AA, Vanbellingen T, Burgunder J, et al. Development and evaluation of maze-like puzzle games to assess cognitive and motor function in aging and neurodegenerative diseases. Front Aging Neurosci 2020 Apr 21;12:87 [FREE Full text] [doi: 10.3389/fnagi.2020.00087] [Medline: 32372942]
40. Chesham A, Gerber SM, Schütz N, Saner H, Gutbrod K, Müri RM, et al. Search and match task: development of a Taskified match-3 puzzle game to assess and practice visual search. JMIR Serious Games 2019 May 09;7(2):e13620 [FREE Full text] [doi: 10.2196/13620] [Medline: 31094325]
41. Manera V, Petit P, Derreumaux A, Orvieto I, Romagnoli M, Lyttle G, et al. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. Front Aging Neurosci 2015 Mar 17;7:24 [FREE Full text] [doi: 10.3389/fnagi.2015.00024] [Medline: 25852542]
42. Vallejo V, Wyss P, Rampa L, Mitache AV, Müri RM, Mosimann UP, et al. Evaluation of a novel serious game based assessment tool for patients with Alzheimer's disease. PLoS One 2017 May 4;12(5):e0175999 [FREE Full text] [doi: 10.1371/journal.pone.0175999] [Medline: 28472049]
43. Wang CS, Wu J, Hsu W, Chien P, Chen P, Huang Y, et al. Using self-administered game-based cognitive assessment to screen for degenerative dementia: a pilot study. J Alzheimers Dis 2022;86(2):877-890. [doi: 10.3233/JAD-215142] [Medline: 35147533]
44. Manning V, Piercy H, Garfield JB, Clark SG, Andrabi MN, Lubman DI. A personalized approach bias modification smartphone app ("SWiPE") to reduce alcohol use: open-label feasibility, acceptability, and preliminary effectiveness study. JMIR Mhealth Uhealth 2021 Dec 10;9(12):e31353 [FREE Full text] [doi: 10.2196/31353] [Medline: 34890355]
45. McWilliams EC, Barbey FM, Dyer JF, Islam MN, McGuinness B, Murphy B, et al. Feasibility of repeated assessment of cognitive function in older adults using a wireless, mobile, Dry-EEG headset and tablet-based games. Front Psychiatry 2021 Jun 25;12:574482 [FREE Full text] [doi: 10.3389/fpsyt.2021.574482] [Medline: 34276428]
46. Palmisano S, Allison RS, Kim J. Cybersickness in head-mounted displays is caused by differences in the user's virtual and physical head pose. Front Virtual Real 2020 Nov 12;1:66. [doi: 10.3389/frvir.2020.587698]
47. Paliokas I, Kalamaras E, Votis K, Doumpoulakis S, Lakka E, Kotsani M, et al. Using a virtual reality serious game to assess the performance of older adults with frailty. Adv Exp Med Biol 2020;1196:127-139. [doi: 10.1007/978-3-030-32637-1_13] [Medline: 32468314]
48. Zygouris S, Giakoumis D, Votis K, Doumpoulakis S, Ntovas K, Segkouli S, et al. Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. J Alzheimers Dis 2015;44(4):1333-1347. [doi: 10.3233/JAD-141260] [Medline: 25428251]
49. Zucchella C, Sinforiani E, Tassorelli C, Cavallini E, Tost-Pardell D, Grau S, et al. Serious games for screening pre-dementia conditions: from virtuality to reality? A pilot project. Funct Neurol 2014;29(3):153-158 [FREE Full text] [Medline: 25473734]
50. Boot WR, Champion M, Blakely DP, Wright T, Souders DJ, Charness N. Video games as a means to reduce age-related cognitive decline: attitudes, compliance, and effectiveness. Front Psychol 2013;4:31 [FREE Full text] [doi: 10.3389/fpsyg.2013.00031] [Medline: 23378841]
51. Lee Y, Kozar KA, Larsen KR. The technology acceptance model: past, present, and future. Commun Assoc Inf Syst 2003;12:50. [doi: 10.17705/1cais.01250]
52. Rivas Costa C, Fernández Iglesias MJ, Anido Rifón LE, Gómez Carballa M, Valladares Rodríguez S. The acceptability of TV-based game platforms as an instrument to support the cognitive evaluation of senior adults at home. PeerJ 2017;5:e2845 [FREE Full text] [doi: 10.7717/peerj.2845] [Medline: 28070464]
53. Gourlan M, Sarrazin P, Trouilloud D. Motivational interviewing as a way to promote physical activity in obese adolescents: a randomised-controlled trial using self-determination theory as an explanatory framework. Psychol Health 2013 Nov;28(11):1265-1286 [FREE Full text] [doi: 10.1080/08870446.2013.800518] [Medline: 23756082]
54. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
55. Gledhill JA, Rodary C, Mahé C, Laizet C. [French validation of the revised Piper Fatigue Scale]. Rech Soins Infirm 2002 Mar(68):50-65. [Medline: 12001628]

56. Bottiroli S, Cavallini E. Can computer familiarity regulate the benefits of computer-based memory training in normal aging? A study with an Italian sample of older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2009 Jul 29;16(4):401-418. [doi: 10.1080/13825580802691763] [Medline: 19253069]
57. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:189-194.
58. Brooke J. SUS: a retrospective. J Usability Stud 2013;8:29-40 [FREE Full text]
59. Bangor AP, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4:114-123 [FREE Full text]
60. Joddrell P, Astell AJ. Implementing accessibility settings in touchscreen apps for people living with dementia. Gerontology 2019 Apr 24;65(5):560-570. [doi: 10.1159/000498885] [Medline: 31018201]
61. Przybylski AK, Rigby CS, Ryan RM. A motivational model of video game engagement. Rev Gen Psychol 2010 Jun 01;14(2):154-166. [doi: 10.1037/a0019440]
62. Uysal A, Yildirim IG. Self-determination theory in digital games. In: Bostan B, editor. Gamer Psychology and Behavior. Cham, Switzerland: Springer; 2016:123-135.
63. Vaz de Carvalho C. Dynamic serious games balancing. In: Proceedings of the 5th International Conference Serious Games, Interaction, and Simulation. 2016 Presented at: SGAMES '15; September 16-18, 2015; Novedrate, Italy p. 33 URL: https://eudl.eu/doi/10.1007/978-3-319-29060-7_5 [doi: 10.1007/978-3-319-29060-7_5]
64. Plechatá A, Sahula V, Fayette D, Fajnerová I. Age-related differences with immersive and non-immersive virtual reality in memory assessment. Front Psychol 2019 Jun 11;10:1330 [FREE Full text] [doi: 10.3389/fpsyg.2019.01330] [Medline: 31244729]

Abbreviations

AD: Alzheimer's disease
CVLT: California Verbal Learning Test
GBCA: game-based cognitive assessment
ISO: International Organization for Standardization
MCI: mild cognitive impairment
NL: Numberlink
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SASG: Smart Aging Serious Game
SMT: search and match task
SUS: system usability scale
VST: Virtual Supermarket Test

Edited by R Yang; submitted 09.08.24; peer-reviewed by M Friehs, N Vita-Barrull, A Katz; comments to author 02.01.25; revised version received 11.02.25; accepted 29.03.25; published 11.06.25.

Please cite as:
Mantell R, Hwang YIJ, Dark M, Radford K, Kasumovic M, Monds L, Schofield PW, Butler T, Withall A
Evaluating the User Experience and Usability of Game-Based Cognitive Assessments for Older People: Systematic Review
JMIR Aging 2025;8:e65252
URL: https://aging.jmir.org/2025/1/e65252
doi:10.2196/65252
PMID:

The Impact of Social Stress and Healthy Lifestyle on the Mortality of Chinese Older Adults: Prospective Cohort Study

Jin Yang[1*], MD, PhD; Jilong Huang[1*], MPH; Qingmei Huang[1], MD, PhD; Jian Gao[1], MD, PhD; Dan Liu[1], MD, PhD; Zhihao Li[1], MD, PhD; Yuebin Lv[2], MD, PhD; Xiaoming Shi[2], MD, PhD; Chen Mao[1,3], MD, PhD

[1]Department of Epidemiology, School of Public Health, Southern Medical University, No. 1023-1063, Shatai South Road, Baiyun District, Guangzhou, Guangdong, China

[2]China CDC Key Laboratory of Environment and Population Health, National Institute of Environmental Health,, Chinese Center for Disease Control and Prevention, Beijing, China

[3]National Institute of Health Data Science of China, Southern Medical University, Guangzhou, Guangdong, China

[*]these authors contributed equally

Corresponding Author:
Chen Mao, MD, PhD
Department of Epidemiology, School of Public Health, Southern Medical University, No. 1023-1063, Shatai South Road, Baiyun District, Guangzhou, Guangdong, China

Abstract

Background: With social progress, social stress (SS) has become a key factor affecting health. Unhealthy lifestyles may exacerbate these effects. However, the relationship between SS, lifestyle, and older adults' mortality rate still needs to be studied.

Objective: This study aimed to explore the relationship between SS and all-cause mortality in Chinese older adults, as well as the influence of healthy lifestyle factors.

Methods: Three groups of SS were defined through latent class analysis: low, medium, and high. We created a healthy lifestyle index based on smoking, alcohol consumption, physical activity, and diet. Multivariable Cox proportional hazards models, interaction analyses, and mediation analyses were conducted.

Results: The Chinese Longitudinal Healthy Longevity Survey (CLHLS) datasets included participants from 806 cities and counties across 23 provinces in China from 1998 to 2018. In this study, participants were recruited from 4 waves of the CLHLS (2005, 2008, 2011, and 2014). Finally, 19,236 participants were included in this study, of which 6891 (35.8%) had low SS, 11,662 (60.6%) had medium SS, and 683 (3.6%) had high SS. In the fully adjusted model, the hazard ratio (HR) for medium SS was 1.16 (95% CI 1.11 - 1.20; *P*<.001), and for high SS, it was 1.28 (95% CI 1.18 - 1.40; *P*<.001) compared to the low SS group. For individuals aged ≥80 years, the medium SS group had a 28% (HR 1.28, 95% CI 1.22 - 1.34; *P*<.001) increased mortality risk, and the high SS group had a 38% (HR 1.38, 95%CI 1.26 - 1.52; *P*<.001) increased risk compared to the low SS group. Approximately 7% of the association between SS and mortality was mediated through the healthy lifestyle. Under different SS, the lower the healthy lifestyle score, the higher the risk of mortality.

Conclusions: SS was an independent predictor of all-cause mortality in Chinese older adults. The healthy lifestyle mediated this effect to some extent. Unhealthy lifestyle behaviors were associated with a higher risk of mortality at all SS levels.

(JMIR Aging 2025;8:e75942) doi:10.2196/75942

KEYWORDS

social stress; healthy lifestyle; mortality; older adults; cohort study

Introduction

As social progress advances, increasing attention has been given to the impact of chronic social stress (SS) on human health. SS is a complex and multifaceted construct, typically encompassing 3 key areas, such as interpersonal relationships, economic conditions, and the living environment. Interpersonal stress arises from demands and expectations within family, friendships, and work relationships [1]. Economic stress is primarily linked to concerns over income or job stability [2] while environmental stress involves factors, such as housing stress and access to social services [3]. All of these factors can contribute to an individual's overall level of SS. Clinical and biological studies have demonstrated that long-term exposure to stress can disrupt the normal functioning of biological systems, leading to negative health outcomes, such as cardiovascular disease [4], gastrointestinal disorders [5], and immune dysfunction [6]. In addition, unhealthy lifestyle behaviors, such as poor diet, physical inactivity, and smoking, are well-documented risk

factors for a range of chronic diseases and premature mortality [7], and may further compound the adverse effects of SS.

China is undergoing a significant demographic shift, with a rapidly aging population. As such, the health and well-being of older adults have become key areas of focus. While stress affects individuals at all stages of life, its impact on older adults warrants particular attention. Studies have shown that stress may be linked to poorer health in older adults [8-10]. However, more research is needed to understand how SS affects the health of older adults and the impact of lifestyle.

This study aimed to investigate the impact of different levels of SS on mortality among Chinese adults older than 65 years, using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). In addition, the study investigated the possible impact of a healthy lifestyle on the relationship between SS and all-cause mortality.

Methods

Study Population

The data for this study were drawn from the Chinese Longitudinal Healthy Longevity Survey (CLHLS), a prospective cohort study described in detail in previous publications [11,12]. In brief, the CLHLS was based on a dynamic, multistage stratified sampling design, which includes participants from 806 cities and counties across 23 provinces in China from 1998 to 2018. The survey covers 85% of the Chinese population, with a focus on individuals aged 65 years and older [13]. Data collected include information on demographic characteristics, socioeconomic status, lifestyle factors, and health-related variables. For this analysis, 27,585 participants were recruited from 4 waves of the CLHLS (2005, 2008, 2011, and 2014). During the follow-up surveys, the survival status of the participants was ascertained as surviving, dead, or lost to follow-up. Participants were excluded if they met any of the following criteria: (1) lost to follow-up or dead before (n=5285), (2) age younger than 65 years at baseline (n=201), and (3) missing baseline information (n=2863). After applying these exclusions, the final sample consisted of 19,236 participants.

Assessment of SS

Existing research primarily examines the impact of SS on health from 3 aspects, namely, interpersonal relationships, economic conditions, and housing conditions [14-16]. Accordingly, our study assessed SS among the older adult population in the CLHLS based on these 3 dimensions. Based on the investigation of relationships with marriage, the health of children, and social support, these questions further transformed into 3 questions, such as “Do you live with your spouse?,” “Are your children still alive?,” and “Do you have someone with whom you can talk, confide in, or seek help when needed?” Economic status was measured by the question, “Is all of the financial support sufficient to pay for daily expenses?” Housing stress was evaluated using the question, “How many people are living with you?” Given the living arrangements typical for older adults in China, those living alone or with more than 5 people were categorized as experiencing housing stress (Table S1 in Multimedia Appendix 1).

Assessment of Healthy Lifestyle Score

A healthy lifestyle was defined by 4 factors, such as smoking, alcohol consumption, physical activity, and diet. Participants who had never smoked or had quit smoking for at least 30 years were assigned 1 point [17]. Men who consume less than 62 g of alcohol per day or women who consume less than 41 g per day were assigned 1 point [18]. Physical activities included regular exercise (both aerobic and anaerobic), housework, tasks, outdoor activities, gardening, pet care, reading, playing cards or mahjong, watching television or listening to the radio, and attending social events. The frequency of “almost every day” was scored as 2, “occasionally” was scored as 1, and “rarely or never” was scored as 0. A total physical activity score was calculated by summing the points across these categories and then standardized to a range of 0-1. A score below 0.6 was assigned 0 points while a score of 0.6 or above was assigned 1 point (Table S2 in Multimedia Appendix 1). The dietary score was determined using a standardized food frequency questionnaire that includes fresh vegetables, fresh fruits, legumes, meat, eggs, fish and seafood, salty vegetables, tea, and garlic. Similar to physical activity, the score for each food group was based on frequency of consumption, and the total dietary score is calculated accordingly (Table S3 in Multimedia Appendix 1). The final lifestyle score ranged from 0 to 4, with higher scores indicating a healthier lifestyle [19].

Outcome

Death was confirmed either by official death certificates or through information provided by the participant’s next of kin, local physician, or resident physician. The follow-up period was defined as the time from study enrollment to the date of death or the last follow-up, whichever came first. Loss to follow-up was defined as the inability to contact the participant after at least 3 documented attempts.

Covariates

Several variables were considered as potential confounding factors. The first set included demographic variables, namely, age (y), sex (male or female), residence (urban or rural), education, and occupation prior to retirement. Considering that the participants included in this study were 65 years old and above, 0 years of education was defined as low, less than 5 years of education was defined as medium, and 5 years of education or more was defined as high. Occupations before retirement were divided into farming, employed, and unemployed. The second set included the history of chronic diseases, such as hypertension (yes or no), diabetes (yes or no), heart disease (yes or no), stroke and cerebrovascular disease (yes or no), and cancer (yes or no). The history of chronic diseases was measured with the following question: “Have you been diagnosed by a doctor with the conditions listed below?”

Statistical Analysis Methods of the Results

Continuous variables were summarized as means and SDs while categorical variables were presented as counts (N) and percentages (%). Differences between groups were assessed using analysis of variance for continuous variables and the chi-square test for categorical variables. Cox proportional hazard regression models were used to estimate the hazard ratios (HRs)

and 95% CI for mortality associated with SS and healthy lifestyle scores. The proportional hazards assumption was tested using Schoenfeld residuals. Model 1 was unadjusted; Model 2 adjusted for age and sex; Model 3 further included education level, occupation, and residence; and Model 4 additionally adjusted for healthy lifestyle factors. Mediation analysis was used to further explore the role of healthy lifestyle in the relationship between SS and all-cause mortality. In addition, the product term of SS and healthy lifestyle score was added to the model to examine their multiplicative interaction. The relative excess risk due to interaction and its corresponding 95% CI were used to assess the interaction on the additive scale. This was calculated using the coefficients and SEs of the product term, SS, and healthy lifestyle score, along with the covariance matrix.

Latent class analysis (LCA) was used to classify participants into distinct SS groups. Convergence was defined as a maximum absolute deviation of 0.000001 between parameter estimates across successive iterations. Iterations terminated when this deviation threshold was reached. Model selection was guided by criteria, including Akaike Information Criterion, Bayesian Information Criterion, likelihood ratio statistic (G^2), and entropy value. Entropy, ranging from 0 to 1, assesses the accuracy of classification in LCA, with higher values indicating better model classification accuracy. In this study, after convergence failed for the 6-class model, 5 latent class groups were selected for comparison, with optimal model performance observed at 3 latent classes based on the highest entropy values (Tables S4-S6 in Multimedia Appendix 1).

Furthermore, when we analyze the impact of a healthy lifestyle on mortality under different SS conditions, the small number of healthy lifestyle groups in the high SS group may affect the results. Therefore, we combined the medium SS group and the high SS group into the medium-high SS group for analysis. We conducted a subgroup analysis stratified by age, sex, education, residence, and occupation to present the differences. We performed several sensitivity analyses to test the robustness of our results: (1) Excluded deaths that occurred within the first 2 years of follow-up. (2) Excluded the mediating effect of each lifestyle factor on the relationship between SS and all-cause mortality. (3) The results were evaluated again by weighting the lifestyle scores. (4) Due to the low proportion of participants with lifestyle scores of 2 and 3, they were combined into 1 group for analysis.

All analyses were based on R version 4.3.2 (R Foundation for Statistical Computing). All statistical tests were 2-sided, with $P<.05$ considered to be indicative of statistical significance.

Ethical Considerations

This study was approved by the Ethics Committee of Peking University (IRB00001052 −13074), and informed consent was obtained from all participants. All data were anonymized and deidentified to protect participant privacy. Participants provided written consent for the publication of their anonymized data and quotes in this study.

Results

Baseline Characteristics

A total of 19,236 Chinese adults aged 65 years and older were included in this prospective cohort study (Figure 1). Among them, 6891 (35.8%) people had low SS, 11,662 (60.6%) people had medium SS, and 683 (3.6%) people had high SS. The overall mean age of the participants was 88 (SD 12) years. Overall, females, rural residents, medium education, and agriculture make up a larger proportion of the total population. The high SS group exhibited a higher proportion of individuals experiencing several adverse conditions, including the loss of children, lack of social support, living without a spouse, insufficient financial resources to cover daily expenses, and housing stress. Overall, 6827 (35.5%) participants had a healthy lifestyle score of 0, 7017 (36.5%) participants scored 1, 2633 (13.7%) participants scored 2, and 2759 (14.3%) participants scored ≥3. Chronic disease conditions and other details can be seen in Table 1.

Figure 1. Flowchart of participant selection. CLHLS: Chinese Longitudinal Healthy Longevity Survey.

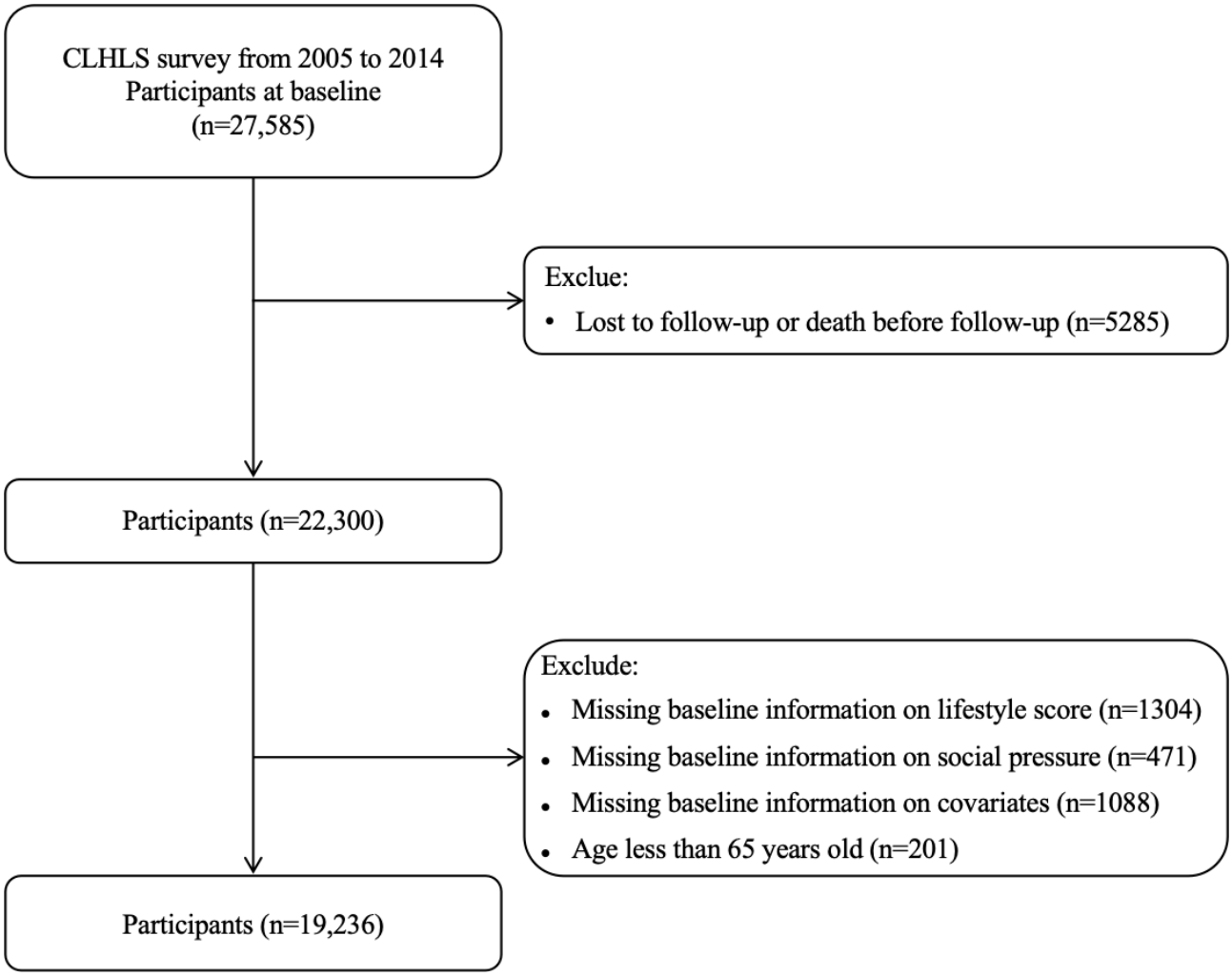


Table . Baseline characteristics of the study participants.

Variables	Overall (N=19,236)	Social stress		
		Low (n=6891)	Medium (n=11,662)	High (n=683)
Age (year), mean (SD)	88 (12)	81 (11)	91 (10)	91 (10)
Sex (female), n (%)	11195 (58.2)	2792 (40.5)	7964 (68.3)	439 (64.3)
Residence (urban), n (%)	2338 (12.2)	908 (13.2)	1368 (11.7)	62 (9.1)
Education[a], n (%)				
High	3436 (17.9)	1950 (28.3)	1429 (12.3)	57 (8.3)
Medium	12487 (64.9)	3432 (49.8)	8513 (73.0)	542 (79.4)
Low	3313 (17.2)	1509 (21.9)	1720 (14.7)	84 (12.3)
Occupation, n (%)				
Agriculture	13042 (67.8)	4382 (63.6)	8140 (69.8)	520 (76.1)
Unemployment	2560 (13.3)	671 (9.7)	1785 (15.3)	104 (15.2)
Employment	3634 (18.9)	1838 (26.7)	1737 (14.9)	59 (8.6)
Are your children alive? (Yes), n (%)	18519 (96.3)	6810 (98.8)	11232 (96.3)	477 (69.8)
Do you have someone with whom you can talk, confide in, or seek help when needed? (Yes), n (%)	17676 (91.9)	6750 (98.0)	10778 (92.4)	148 (21.7)
Do you live with your spouse? (Yes), n (%)	5637 (29.3)	5616 (81.5)	0 (0.0)	19 (3.1)
Is all of the financial support sufficient to pay for daily expenses? (Yes), n (%)	14914 (77.5)	4381 (63.6)	10532 (90.3)	1 (0.1)
Are you living alone or with ≥5 people? (No), n (%)	11141 (57.9)	5545 (80.5)	5333 (45.7)	263 (38.5)
Healthy lifestyle score[b], n (%)				
≥3	2759 (14.3)	1190 (17.3)	1540 (13.2)	29 (4.2)
2	2633 (13.7)	1040 (15.1)	1460 (12.5)	133 (19.5)
1	7017 (36.5)	2234 (32.4)	4409 (37.8)	374 (54.8)
0	6827 (35.5)	2427 (35.2)	4253 (36.5)	147 (21.5)
Hypertension (yes), n (%)	4981 (25.9)	2032 (29.5)	2829 (24.3)	120 (17.6)
Diabetes (yes), n (%)	1964 (10.2)	824 (12.0)	1107 (9.5)	33 (4.8)
Heart disease (yes), n (%)	2935 (15.3)	1214 (17.6)	1657 (14.2)	64 (9.4)
Stroke and cerebrovascular disease (yes), n (%)	2158 (11.2)	898 (13.0)	1206 (10.3)	54 (7.9)
Cancer (yes), n (%)	1042 (5.4)	404 (5.9)	618 (5.3)	20 (2.9)
Mortality, n (%)	13999 (72.8)	3986 (57.8)	9425 (80.8)	588 (86.1)

[a]Education: low (0 y), medium (<5 y), high (≥5 y).

[b]Healthy lifestyle was defined by 4 factors: smoking, alcohol consumption, physical activity, and diet. Healthy lifestyle score ranges from 0 to 4, with higher scores indicating a healthier lifestyle.

Association of SS and All-Cause Mortality

In the unadjusted analysis (Model 1), compared with the low SS group, the results showed the HRs of the medium and high SS groups were 2.05 (95% CI 1.98 - 2.13) and 2.47 (95% CI 2.26 - 2.69). After adjusting for age and sex (Model 2), the HRs of the medium and high SS groups were 1.15 (95% CI 1.11 - 1.20) and 1.37 (95% CI 1.26 - 1.50). In Model 3, which further adjusted for local education, occupation, and healthy lifestyle score based on Model 2, the HRs of the medium and high SS groups were 1.15 (95% CI 1.11 - 1.20) and 1.28 (95% CI 1.17 - 1.39). In the fully adjusted Model 4, the mortality risk for medium SS was 1.16 (95% CI 1.11 - 1.20), and for high

SS, it was 1.28 (95% CI 1.18 - 1.40) compared to the low SS groups (Table 2).

Table . Association of social stress and all-cause mortality.

Social stress	Model 1[a] HR[b] (95% CI)	Model 2[c] HR (95% CI)	Model 3[d] HR (95% CI)	Model 4[e] HR (95% CI)	*P* value
Low[f]	1 (Reference)	1 (Reference)	1 (Reference)	1 (Reference)	—[g]
Medium	2.05 (1.98-2.13)	1.15 (1.11-1.20)	1.15 (1.11-1.20)	1.16 (1.11-1.20)	<.001
High	2.47 (2.26-2.69)	1.37 (1.26-1.50)	1.28 (1.17-1.39)	1.28 (1.18-1.40)	<.001

[a]Model 1: unadjusted.

[b]Hazard ratio

[c]Model 2: adjusted for age and sex.

[d]Model 3: adjusted for residence, education, occupation, and healthy lifestyle score based on Model 2.

[e]Model 4: adjusted for hypertension, diabetes, heart disease, cancer, stroke, and cerebrovascular disease based on Model 3.

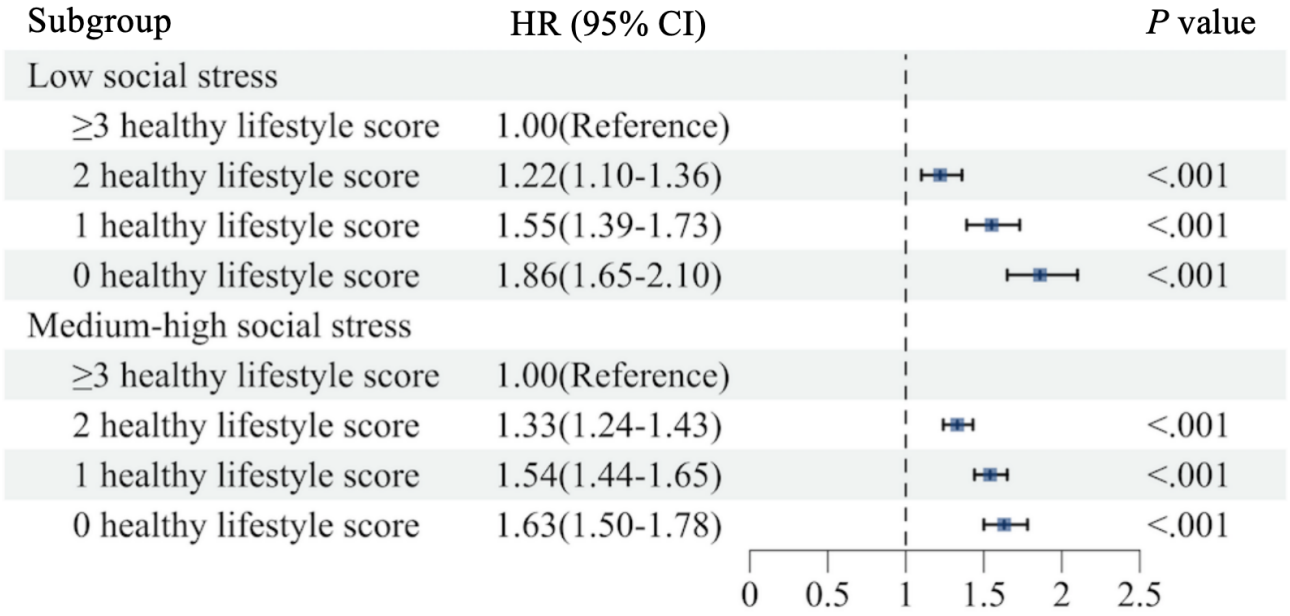
[f]The group is set as a reference group.

[g]Not applicable.

Impact of a Healthy Lifestyle on All-Cause Mortality Across Different SS Groups

The lower the healthy lifestyle score, the higher the risk of death compared with a healthy lifestyle score ≥3 (Table S7 in Multimedia Appendix 1). For the low SS group, individuals with a higher healthy lifestyle score had a lower mortality risk. Specifically, compared to those with a healthy lifestyle score of 3 or more, individuals with a score of 2 had a 22% increased risk of mortality (HR 1.22, 95% CI: 1.10 - 1.36), those with the score of 1 had a 55% increased risk (HR 1.55, 95% CI 1.39 - 1.73), and those with the score of 0 had an 86% increased risk (HR 1.86, 95% CI 1.65 - 2.10). Further combining the medium and high SS groups, the results found that individuals with a score of 2 had a 33% higher mortality risk (HR 1.33, 95% CI 1.24 - 1.43) compared to those with a score of 3 or more. The score of 1 was associated with a 54% higher risk (HR 1.54, 95% CI 1.44 - 1.65), and the score of 0 was associated with a 63% higher risk (HR 1.63, 95% CI 1.50 - 1.78) (Figure 2).

Figure 2. Impact of a healthy lifestyle on all-cause mortality across different social stress groups. All covariates were adjusted. HR: hazard ratio.



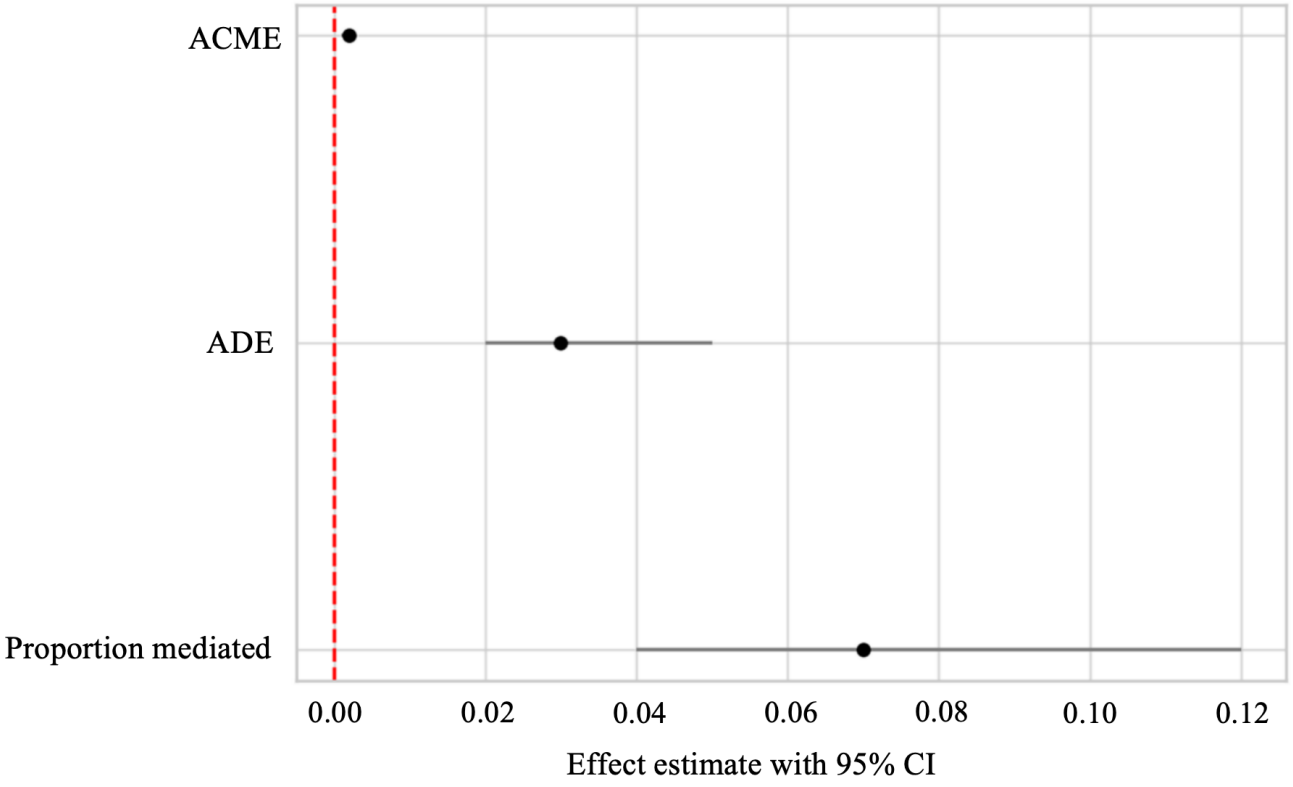
Mediation and Interaction Analysis of a Healthy Lifestyle on Associations of SS With All-Cause Mortality

The proportion of the total effect of SS on all-cause mortality that is mediated through a healthy lifestyle was 0.07 (95% CI 0.04-0.12), with a *P* value <.001. This suggests that approximately 7% of the association between SS and mortality is explained by the mediation effect of a healthy lifestyle. The average direct effect of SS on all-cause mortality, independent of the mediation through a healthy lifestyle, was 0.03 (95% CI 0.02-0.05), with a *P* value <.001. The average causal mediation effect was 0.002 (95% CI 0.001-0.002), with a *P* value <.001. These findings demonstrate that a healthy lifestyle partially mediates the relationship between SS and all-cause mortality (Figure 3).

After multiplicative interaction, the interaction term exponentiated coefficient (exp [coef]) is 0.941, and the *P* value

was less than .001, indicating that there was a significant multiplicative interaction between social pressure and a healthy lifestyle. The additive interaction results showed that the relative excess risk due to interaction was measured to be 0.127 (95% CI 0.013 - 0.240), indicating the existence of a positive additive interaction. The attributable proportion due to interaction was found to be 0.071 (95% CI 0.017 - 0.124), indicating that 7.1% of the combined mortality risk caused by social pressure and unhealthy lifestyle can be attributed to the synergistic effect of both factors. The Synergy Index was calculated to be 1.189 (95% CI 1.060 - 1.334), indicating that when SS coexists with unhealthy lifestyles, the risk of death is 1.189 times higher than the sum of their individual effects.

Figure 3. The mediating effect of a healthy lifestyle in the relationship between social stress and all-cause mortality. ACME: average causal mediation effects; ADE: average direct effects.



Stratified and Sensitivity Analyses

We analyzed the effects of SS and a healthy lifestyle on the oldest- and nonoldest-old. The results showed that there was a significant association between SS and mortality risk in both the ≥80 years and <80 years age groups (Table S8 in Multimedia Appendix 1). In the older adults and nonolder adults groups, regardless of stress level, the mortality risk (HR) increased significantly as the healthy lifestyle score decreased (from ≥3 points to 0 points) (Figure S1 and Table S9 in Multimedia Appendix 1). In the older adults and nonolder adults groups, healthy lifestyle partially mediated the relationship between SS and mortality risk (Table S10 in Multimedia Appendix 1). The associations of SS, healthy lifestyle, and all-cause mortality were generally similar across sex (Table S11 in Multimedia Appendix 1), residence (Table S12 in Multimedia Appendix 1), occupation (Table S13 in Multimedia Appendix 1), and education (Table S14 in Multimedia Appendix 1). The associations of SS, healthy lifestyle, and all-cause mortality did not change appreciably when we excluded participants who died within 1 or 2 years of follow-up (Table S15 in Multimedia Appendix 1). The mediating effects of different components of a healthy lifestyle (smoking, alcohol consumption, physical activity, and diet scores) on the relationship between SS and mortality were further analyzed (Table S16 in Multimedia Appendix 1). Association of SS with all-cause mortality after adjustment for weighted lifestyle score and other covariates (Table S17 in Multimedia Appendix 1). Since the population with lifestyle score=2 and lifestyle score=3 each accounted for 1/3 of the other population (lifestyle score=0 or 1), we combined the population with score 2 and score 3 for sensitivity analysis, and the results remained robust (Table S18 in Multimedia Appendix 1).

Discussion

Principal Findings

In this large prospective cohort study, we investigated the impact of SS and a healthy lifestyle on mortality among older adults in China. Our findings showed that both medium and high levels of SS were associated with increased mortality risks, with HRs of 1.16 and 1.28, respectively, compared to the low SS group after full adjustment for covariates. For individuals aged <80 years and ≥80 years, the HRs of the high SS group were 2.09 and 1.38. SS has a greater impact on the mortality of older adults younger than 80 years. These results underscore the significant impact of SS on health outcomes in older adults. Approximately 7% of the association between SS and mortality was mediated through the healthy lifestyle. Under different SS, the lower the healthy lifestyle score, the higher the risk of mortality. These

results provide valuable insights into the complex interplay between SS and health behaviors in shaping mortality risks among older adults in China.

Chronic SS has been shown to induce a range of physiological and behavioral changes that contribute to the development or progression of various health conditions [20]. Previous studies have demonstrated that prolonged exposure to stress elevates plasma cortisol levels, which may lead to memory impairment [21]. Mild stress could help improve cognitive function, but if the intensity of stress exceeded a predetermined threshold (which is different for each person), it could lead to cognitive impairment [22]. Stress could affect the function of the immune system by regulating the processes of the central nervous system and the neuroendocrine system [23]. In addition, stress could affect the occurrence of cardiovascular disease, gastrointestinal disease, endocrine disease, and other diseases [24-26]. A study from China explored the relationship between SS and health inequality and confirmed that SS has a greater impact on people with low economic status [14]. However, most existing studies focus on the impact of SS on psychological or psychiatric diseases. In addition, there is currently a lack of research on the impact of SS on the survival of Chinese older adults and the role of lifestyle in this association. With the aging trend in China, it is necessary to explore the relationship between SS and the health of older adults [27,28]. In our study, we found that compared with low SS, moderate and high SS increased the risk of death in Chinese people aged 65 years and more than 65 years to varying degrees, and lifestyle played a partial mediating role. This study provides causal support for the relationship between SS and the health of older adults through prospective analysis.

Although stress and its effects occur at all stages of an individual's life, with the rapid development of China's aging population, the social pressure faced by older adults has increased, mainly including pressure from interpersonal relationships, the economy, and the environment [29,30]. These components are interrelated and should be examined collectively in relation to health outcomes in older adults. Our study divided these 3 variables into 3 levels, such as low, medium, and high, based on LCA analysis. The results showed that among adults aged 65 years and older, compared with low SS, medium SS increased the risk of death by 16%, and high SS increased the risk of death by 28%. Approximately 7% of the association between SS and mortality is explained by the mediation effect of a healthy lifestyle. This indicates that SS directly influences mortality risk among older adults, and although a healthy lifestyle partially mediates this relationship, the effect size is small. We speculated that there may be several reasons. First, for older adults, the impact of SS may have already had a cumulative effect on their physical health earlier in life, and adopting a healthy lifestyle may not be sufficient to fully offset this long-term negative impact. Second, with age, declining physical functioning, accumulated chronic diseases, and a weakened immune system may reduce older adults' ability to benefit from a healthy lifestyle. Third, social pressure is a complex stressor that may affect the health of older adults' pathways, such as mental well-being, immune function, quality of life, and other mechanisms. Older adults may face higher levels of loneliness and reduced social support, and a healthy lifestyle may not sufficiently mitigate the physiological and psychological effects of these pressures.

Furthermore, participants aged ≥80 years were analyzed separately from younger older adults. We found that SS had a greater impact on mortality risk among older adults aged less than 80 years. Compared with the low SS group, mortality risk increased by 35% and 109% in the medium and high SS groups, respectively. Among adults aged 80 years and older than 80 years, the risk of death increased by 28% and 38% in the medium and high SS groups, respectively, compared to the low SS group. This suggests that SS may pose a greater threat to the survival of younger-old adults. Several factors may help explain this phenomenon. First, younger-old adults may be in a critical stage of lifestyle change, such as facing adaptation after retirement, changes in social roles, etc. Stress may accelerate the occurrence or aggravation of cardiovascular diseases, diabetes, and other chronic diseases, potentially leading to premature death [31]. Second, younger-older adults may not fully cope with SS, resulting in a surge in emotions, such as anxiety and depression [32]. Third, younger-older adults may experience greater economic pressure, including ongoing family expenses and housing challenges [33]. Fourth, in response to a rapidly changing social environment, younger-old adults may adopt unhealthy coping strategies, such as smoking or alcohol consumption, which can exacerbate health risks [34]. Therefore, our findings suggest that age-specific interventions may be necessary to reduce the impact of SS on the health and survival of older adults. This study investigated the interaction between SS and healthy lifestyle on all-cause mortality, using both multiplicative and additive interaction models. The findings revealed a statistically significant interaction effect, suggesting that the relationship between SS and mortality is modified by a healthy lifestyle. The additive interaction highlights the importance of addressing both SS and unhealthy lifestyles simultaneously. Interventions targeting stress reduction alongside lifestyle promotion could yield greater mortality benefits than addressing either factor in isolation.

Studies have shown that SS affects health through a variety of biological pathways. For example, long-term SS may increase cortisol levels in the body, which may lead to decreased immune function and enhanced inflammatory response, thereby increasing the risk of illness or death [35]. Our study suggests that reducing SS among older adults may help improve their quality of life. Moreover, our study found that social pressure had a greater impact on mortality risk among males, as well as among older adults who lived in urban areas, had never been employed, and had received more than 5 years of education. These findings offer valuable insights for future research on SS among older adults.

Our study has several major strengths, including a large sample size and a prospective cohort study. To our knowledge, no previous study has provided a detailed analysis of SS and mortality among the older adult population in China. Nevertheless, our study also has some limitations. First, data related to SS and data related to healthy lifestyles were based on self-reports, which, despite strict collection criteria, are still prone to measurement errors. Second, cause-of-death data were

not comprehensively recorded, limiting our ability to assess the impact of SS on cause-specific mortality among older adults. Third, residual confounding could not be fully eliminated due to limited covariates, such as the absence of genetic data and information on other health conditions. Fourth, this study only focuses on the social pressure composed of interpersonal relationships, economic status, and living environment and may not have considered other related factors. Fifth, although our lifestyle score included key behavioral factors, such as smoking, diet, physical activity, and alcohol consumption, other potentially important components, such as sleep duration and quality, perceived stress, and sedentary behavior, were not included due to data limitations. Sixth, we did not consider the potential changes in lifestyle factors during this period, which may affect the observed results. Therefore, more relevant studies are needed to provide evidence for the present findings.

Conclusions

The results showed that SS was an important risk factor for all-cause mortality in the Chinese older adults. This association remained strong even after adjusting for various sociodemographic and health-related factors, highlighting the importance of SS as a determinant of mortality risk in the Chinese older adults. A healthy lifestyle only partially mediated the relationship between SS and survival among older adults. However, an unhealthy lifestyle was associated with an increased risk of mortality, regardless of SS level.

Acknowledgments

The authors express their gratitude to the participants and staff involved in data collection and management in the Chinese Longitudinal Healthy Longevity Study. This study was supported by the National Natural Science Foundation of China (82425052) and the Postdoctoral Fellowship Program of CPSF (GZC20231052).

Data Availability

Data are available from the Chinese Longitudinal Healthy Longevity Study upon request [36]. Deidentified datasets will be provided following approval, with no restrictions on noncommercial use.

Authors' Contributions

XS, YL, and CM collected the data. JY and JH analyzed and interpreted the data. JY and JH wrote the manuscript. QH, JG, DL, and ZL checked the data and results. CM and XS were co-senior authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data supporting the primary analysis.
[DOCX File, 5723 KB - aging_v8i1e75942_app1.docx]

References

1. Avison WR, Turner RJ. Stressful life events and depressive symptoms: disaggregating the effects of acute stressors and chronic strains. J Health Soc Behav 1988 Sep;29(3):253-264. [Medline: 3241066]
2. Farnworth M, Leiber MJ. Strain theory revisited: economic goals, educational means, and delinquency. Am Sociol Rev 1989 Apr;54(2):263. [doi: 10.2307/2095794]
3. La Gory M, Ritchey FJ, Mullis J. Depression among the homeless. J Health Soc Behav 1990 Mar;31(1):87-102. [Medline: 2313079]
4. Franklin BA, Rusia A, Haskin-Popp C, Tawney A. Chronic stress, exercise and cardiovascular disease: placing the benefits and risks of physical activity into perspectiv. Int J Environ Res Public Health 2021 Sep 21;18(18):9922. [doi: 10.3390/ijerph18189922] [Medline: 34574843]
5. Mangiaterra S, Marker L, Cerquetella M, et al. Chronic stress-related gastroenteric pathology in cheetah: relation between intrinsic and extrinsic factors. Biology (Basel) 2022 Apr 15;11(4):606. [doi: 10.3390/biology11040606] [Medline: 35453805]
6. Fali T, Vallet H, Sauce D. Impact of stress on aged immune system compartments: overview from fundamental to clinical data. Exp Gerontol 2018 May;105:19-26. [doi: 10.1016/j.exger.2018.02.007] [Medline: 29427753]
7. Li Y, Schoufour J, Wang DD, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020 Jan 8;368:l6669. [doi: 10.1136/bmj.l6669] [Medline: 31915124]
8. Tehee M, Buchwald D, Booth-LaForce C, Omidpanah A, Manson SM, Goins RT. Traumatic stress, social support, and health among older American Indians: the native elder care study. J Gerontol B Psychol Sci Soc Sci 2019 Jun 14;74(5):908-917. [doi: 10.1093/geronb/gbx163] [Medline: 29304244]
9. Chokkanathan S. Stressors social support and elder mistreatment. Aging Ment Health 2017 Feb;21(2):125-132. [doi: 10.1080/13607863.2015.1081151] [Medline: 26370609]

10. de Frias CM, Whyne E. Stress on health-related quality of life in older adults: the protective nature of mindfulness. Aging Ment Health 2015;19(3):201-206. [doi: 10.1080/13607863.2014.924090] [Medline: 24940847]
11. Wang J, Chen C, Zhou J, et al. Healthy lifestyle in late-life, longevity genes, and life expectancy among older adults: a 20-year, population-based, prospective cohort study. Lancet Healthy Longev 2023 Oct;4(10):e535-e543. [doi: 10.1016/S2666-7568(23)00140-X] [Medline: 37804845]
12. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2]
13. Chinese Longitudinal Healthy Longevity and Happy Family Study (CLHLS-HF). Peking University Open Research Data. URL: https://opendata.pku.edu.cn/dataverse/CHADS [accessed 2024-11-13]
14. Zhang Y. Social stress and health inequality: an empirical study based on CLHLS [Article in Chinese]. South China Population 2023;38(1):1-9. [doi: 10.3969/j.issn.1004-1613.2023.01.001]
15. Birmingham WC, Holt-Lunstad J. Social aggravation: understanding the complex role of social relationships on stress and health-relevant physiology. Int J Psychophysiol 2018 Sep;131:13-23. [doi: 10.1016/j.ijpsycho.2018.03.023] [Medline: 29626504]
16. Engeln M, Fox ME, Lobo MK. Housing conditions during self-administration determine motivation for cocaine in mice following chronic social defeat stress. Psychopharmacology (Berl) 2021 Jan;238(1):41-54. [doi: 10.1007/s00213-020-05657-y] [Medline: 32914243]
17. Tobacco. World Health Organization. URL: https://www.who.int/health-topics/tobacco#tab=tab_1 [accessed 2024-11-14]
18. International guide for monitoring alcohol consumption and related harm. World Health Organization. URL: https://iris.who.int/handle/10665/66529 [accessed 2024-11-14]
19. Zhang YB, Chen C, Pan XF, et al. Associations of healthy lifestyle and socioeconomic status with mortality and incident cardiovascular disease: two prospective cohort studies. BMJ 2021 Apr 14;373:n604. [doi: 10.1136/bmj.n604] [Medline: 33853828]
20. Yaribeygi H, Panahi Y, Sahraei H, Johnston TP, Sahebkar A. The impact of stress on body function: a review. EXCLI J 2017;16:1057-1072. [doi: 10.17179/excli2017-480] [Medline: 28900385]
21. Lupien SJ, McEwen BS, Gunnar MR, Heim C. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nat Rev Neurosci 2009 Jun;10(6):434-445. [doi: 10.1038/nrn2639] [Medline: 19401723]
22. Sandi C. Stress and cognition. Wiley Interdiscip Rev Cogn Sci 2013 May;4(3):245-261. [doi: 10.1002/wcs.1222] [Medline: 26304203]
23. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull 2004 Jul;130(4):601-630. [doi: 10.1037/0033-2909.130.4.601] [Medline: 15250815]
24. Wu ZS. Epidemiological studies on the relationship between psychosocial factors and cardiovascular disease. Zhongguo Yi Xue Ke Xue Yuan Xue Bao 2001 Feb;23(1):73-77. [Medline: 12905825]
25. Konturek PC, Brzozowski T, Konturek SJ. Stress and the gut: pathophysiology, clinical consequences, diagnostic approach and treatment options. J Physiol Pharmacol 2011 Dec;62(6):591-599. [Medline: 22314561]
26. Charmandari E, Tsigos C, Chrousos G. Endocrinology of the stress response. Annu Rev Physiol 2005;67:259-284. [doi: 10.1146/annurev.physiol.67.040403.120816] [Medline: 15709959]
27. Jin S, Li C, Cao X, Chen C, Ye Z, Liu Z. Association of lifestyle with mortality and the mediating role of aging among older adults in China. Arch Gerontol Geriatr 2022 Jan;98:104559. [doi: 10.1016/j.archger.2021.104559]
28. Chen Y, Sun B, Wang C, Zhang K, Lu W, Huang G. Associations of types of grains and lifestyle with all-cause mortality among Chinese adults aged 65 years or older: a prospective cohort study. J Transl Med 2023;21(1):88. [doi: 10.1186/s12967-023-03927-9]
29. Sharifi S, Babaei Khorzoughi K, Khaledi-Paveh B, Rahmati M. Association of intergenerational relationship and supports with cognitive performance in older adults: a systematic review. Geriatr Nurs (Lond) 2023 Jul;52:146-151. [doi: 10.1016/j.gerinurse.2023.05.014]
30. Fangfang H, Xiao H, Shuai Z, et al. Living environment, built environment and cognitive function among older Chinese adults: results from a cross-sectional study. J Prev Alzheimers Dis 2022 Jan;9(1):126-135. [doi: 10.14283/jpad.2021.59] [Medline: 35098983]
31. Schultz WM, Kelli HM, Lisko JC. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178. [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]
32. Thoits PA. Stress and health: major findings and policy implications. J Health Soc Behav 2010;51 Suppl:S41-S53. [doi: 10.1177/0022146510383499] [Medline: 20943582]
33. More than 1 in 10 people in Japan are aged 80 or over. Here's how its ageing population is reshaping the country. World Economic Forum. URL: https://www.weforum.org/stories/2023/09/elderly-oldest-population-world-japan/ [accessed 2024-12-18]
34. Villiers-Tuthill A, Copley A, McGee H, Morgan K. The relationship of tobacco and alcohol use with ageing self-perceptions in older people in Ireland. BMC Public Health 2016 Jul 22;16:627. [doi: 10.1186/s12889-016-3158-y] [Medline: 27448397]

35. McEwen BS. Protective and damaging effects of stress mediators: central role of the brain. Dialogues Clin Neurosci 2006 Dec 31;8(4):367-381. [doi: 10.31887/DCNS.2006.8.4/bmcewen]
36. The Chinese Longitudinal Healthy Longevity Survey (CLHLS)-Longitudinal Data (1998-2018). Peking University Open Research Data. URL: http://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WBO7LK [accessed 2025-06-05]

Abbreviations

CLHLS: Chinese Longitudinal Healthy Longevity Survey
HR: hazard ratio
LCA: latent class analysis
SS: social stress

Edited by J Job; submitted 13.04.25; peer-reviewed by H Duan, Y Wei; revised version received 11.06.25; accepted 13.06.25; published 12.08.25.

Please cite as:
Yang J, Huang J, Huang Q, Gao J, Liu D, Li Z, Lv Y, Shi X, Mao C
The Impact of Social Stress and Healthy Lifestyle on the Mortality of Chinese Older Adults: Prospective Cohort Study
JMIR Aging 2025;8:e75942
URL: https://aging.jmir.org/2025/1/e75942
doi:10.2196/75942

Edentulousness and the Likelihood of Becoming a Centenarian: Longitudinal Observational Study

Xindi Wei[1*], MD; Longfei Zhuang[2*], PhD; Yuan Li[1], PhD; Junyu Shi[1], PhD; Yijie Yang[1], MD; Hongchang Lai[1], PhD; Beilei Liu[1], PhD

[1]Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, 639 Zhizaoju Road, Shanghai, China

[2]Huishang Dental Clinic, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Beilei Liu, PhD
Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, 639 Zhizaoju Road, Shanghai, China

Abstract

Background: In recent decades, the global life expectancy has risen notably to approximately 73.5 years worldwide, coinciding with a rapid growth in the older adult population, which presents a significant public health challenge in promoting healthy aging and longevity.

Objective: This study aimed to prospectively investigate the link between edentulousness and the likelihood of reaching centenarian status among individuals aged 80 years and older.

Methods: Data from the Chinese Longitudinal Healthy Longevity Survey were analyzed. Logistic regression models were used to assess the relationship between edentulousness and the likelihood of becoming a centenarian. Demographic characteristics, lifestyle habits, and disease histories were adjusted as confounding factors. Several sensitivity analyses, including propensity score matching and 2-year lag analyses, were conducted to further assess the association between edentulousness and the likelihood of becoming a centenarian. The correlation between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian was evaluated as well.

Results: The study included 4239 participants aged 80-100 years. After adjusting for all covariates, the likelihood for becoming a centenarian increased in the nonedentulous group compared to the edentulous group (odds ratio [OR] 1.384, 95% CI 1.093 - 1.751). The relationship persisted after propensity score matching analysis (OR 1.272, 95% CI 1.037 - 1.561). The association remained statistically significant after excluding participants with a follow-up duration of less than 2 years (OR 1.522, 95% CI 1.083 - 2.140; *P*=.02). Furthermore, a significant positive association between the number of natural teeth and the likelihood of becoming a centenarian was found after adjusting for all covariates (OR 1.022, 95% CI 1.002 - 1.042; *P*=.03), which aligned with the main results of the study.

Conclusions: The findings revealed that the presence of natural teeth was linked to an increased probability of becoming a centenarian, underscoring the importance of maintaining oral health even in advanced age.

(JMIR Aging 2025;8:e68444) doi:10.2196/68444

KEYWORDS

public health; edentulous; oral-systemic disease; epidemiology; cohort studies

Introduction

Over the past few decades, the global life expectancy at birth has significantly risen to approximately 73.5 years worldwide and 77.6 years in mainland China [1,2]. Additionally, the proportion of individuals aged 65 years and older is rapidly expanding, constituting 13.5% of the population in China [3]. This increase in life expectancy has coincided with a rapid expansion of the aging population, presenting a substantial public health challenge in fostering healthy aging and longevity. However, focusing solely on traditional determinants of longevity is insufficient, underscoring the importance of identifying other modifiable risk factors.

Oral diseases are widespread globally, affecting over 3.5 billion individuals [4,5], yet their influence on longevity is often overlooked. In addition to sociodemographic, genetic, and lifestyle factors [6], tooth loss—the ultimate outcome of oral diseases and a surrogate for overall health status—is essential for healthy aging as well. The retention of natural teeth over a lifetime is a fundamental aspect of overall well-being [7]. Tooth loss can be a potential risk factor for longevity that is independently related to the onset of disability and mortality in old age [7-9]. For instance, the 6-year mortality rate of

individuals with edentulousness (the lack of teeth) who do not use dentures was significantly higher than that of the individuals with ≥20 teeth [10]. Another study based on Baltimore Longitudinal Study of Aging indicated that being edentulous or having fewer than 20 teeth was independently associated with the mortality of older adults [11]. Centenarians comprise those who successfully age and have good resilience, and the number of centenarians is increasing worldwide [12-14]. Centenarians and their offspring demonstrated better oral health, suggesting the potential relationship between natural teeth retention and longevity [8]. Moreover, possessing ≤20 natural teeth was an independent risk factor for frailty among centenarians [3]. However, there is no cohort study that focuses on the impact of natural teeth retention on achieving centenarian status; thus, leveraging aging-focused, big data analytics to understand the intricate relationship among the older adult population becomes imperative.

The objective of this study was to prospectively investigate the association between edentulousness and the likelihood of becoming a centenarian in individuals aged 80 years and older, using data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS)—an aging-focused, nationally representative cohort of the older Chinese population. These findings have the potential to inform interventions aimed at maximizing life expectancy.

Methods

Study Design and Population

The CLHLS is a comprehensive nationwide study that uses random sampling techniques to select participants from half of the counties and cities in 22 out of the 31 provinces across mainland China, covering approximately 85% of the total population [15]. The CLHLS collected data from 8 waves of surveys carried out in 1998, 2000, 2002, 2005, 2008, 2011, 2014, and 2018. Each survey round involved follow-ups with existing participants and the recruitment of new participants.

This cohort study used the baseline data from 1998, and the mortality follow-up data were from 1998, 2000, 2002, 2005, 2008, 2011, 2014, and 2018. As long as the individuals were aged 80 years or older in 1998, they could potentially live to 100 years in 2018. Thus, 6675 participants (aged ≥80 years and <100 years) who had the potential to age to 100 years or older were first included in this study. The participants who were lost to follow-up or had incomplete information on confounders were then excluded, and 4239 participants were included in the final analyses.

This study adheres to the STROBE (Standards for Reporting of Observational Studies in Epidemiology) guidelines [16].

Exposure Variable

The number of natural teeth of the participants was collected in the survey. Participants were divided into two categories based on the existence of natural teeth. The edentulous group was defined as participants with the complete loss of all dentition, while the nonedentulous group was defined as participants with at least 1 tooth.

Outcome Variable

Interviews were conducted with a close family member of the participants who had passed away between the previous wave's interview and the subsequent survey. The year of death and the age at death were collected. Individuals reaching the age of 100 years were defined as centenarians.

Covariates

Age in 1998, sex (male or female), ethnicity (Han Chinese or other), residence (urban or rural), marital status (currently married and living with a spouse, widowed, separated, divorced, or never married), exercise habits (yes or no), smoking status (nonsmoker, former smoker, or current smoker), alcohol use (nondrinker, former drinker, or current drinker), denture use (yes or no), diabetes (yes or no), hypertension (yes or no), stroke or cardiovascular disease (CVD; yes or no), and cancer (yes or no) were selected as covariates [6]. Denture use included the use of complete dentures and removable partial dentures [17]. All data were obtained by a face-to-face interview via a questionnaire by well-trained interviewers from the local Centers for Disease Control and Prevention.

Statistical Analysis

In the descriptive statistics, median and IQR were used for continuous variables, and frequency distributions were used for categorical variables. Continuous variables were compared by ANOVA tests for variables meeting the assumptions of normal distribution and homogeneity of variance, and by Kruskal-Wallis H test for those not meeting these assumptions. Categorical variables were compared using χ^2 tests. The univariate and multivariate logistic regression models were adopted to evaluate the odd ratios (ORs) and 95% CIs pertaining to the association between edentulousness and the likelihood of becoming a centenarian. To control for confounders, three logistic regression models were constructed to eliminate the influence of covariates. No variables were adjusted in model 1. Model 2 adjusted for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits. Model 3 further incorporated adjustments for denture use, diabetes, hypertension, stroke or CVD, and cancer. Several sensitivity analyses were performed. First, a 1:1 propensity score matching (PSM) analysis was conducted to balance the differences between the edentulous and nonedentulous groups using the *MatchIt* R package [18], which adjusted for the age in 1998, sex, ethnicity, marital status, exercise habits, smoking status, alcohol use, denture use, diabetes, hypertension, stroke or CVD, and cancer. The data after PSM were then analyzed using logistic regression to confirm the association between edentulousness and the likelihood of becoming a centenarian, with only the covariates that were still significant after PSM being adjusted. Second, a 2-year lag analysis was conducted by excluding individuals whose follow-ups were less than 2 years. Third, the association between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian was explored by logistic regression models.

A 2-sided *P* value <.05 was considered statistically significant in all analyses. All statistical analyses were performed by R (version 4.3.1; R Foundation for Statistical Computing).

Ethical Considerations

Ethical approval was granted by the Biomedical Ethics Committee of Peking University (IRB00001052-13,074), and all participants or their proxy respondents provided informed consent without receiving financial compensation. The data have been anonymized to protect the privacy of participants.

Results

Baseline Characteristics of the Study Sample

As depicted in Figure 1, this study included 4239 participants from the CLHLS 1998‐2018. Table 1 demonstrated the baseline characteristics of these participants. Significant differences were detected among the participants in various factors, including age (P<.001), sex (P<.001), ethnicity (P=.03), denture use (P<.001), smoking status (P<.001), alcohol use (P<.001), and hypertension (P=.02). There were no significant differences between the two groups in terms of residence (P=.09), exercise habits (P=.87), diabetes (P>.99), stroke or CVD (P=.17), and cancer (P=.47). Overall, 607 (14.3%) of the 4239 individuals became centenarians during follow-up, and 264 (43.5%) of these 607 centenarians were edentulous.

Figure 1. Flowchart of the process for the selection of eligible participants. CLHLS: Chinese Longitudinal Healthy Longevity Survey.

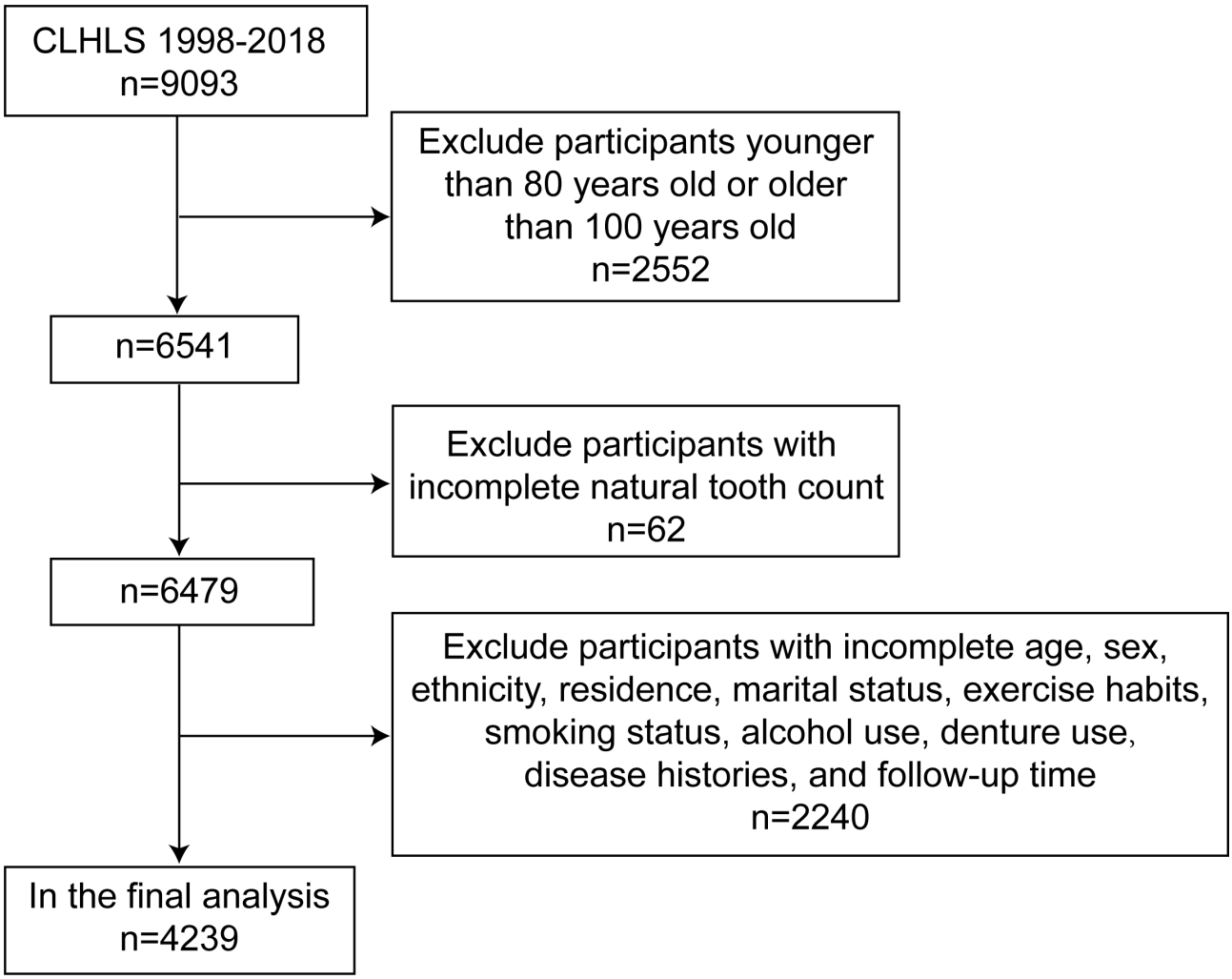


Table . Characteristics of the participants (n=4239).

Characteristics	Noncentenarians (n=3632)	Centenarians (n=607)	Statistic (*df*)	*P* value
Age (years), median (IQR)	88 (84-93)	98 (96-99)	1056.900 (1)	<.001[a]
Sex, n (%)			40.334 (1)	<.001[b]
Male	1795 (49.4)	215 (35.4)		
Female	1837 (50.6)	392 (64.6)		
Ethnicity, n (%)			4.596 (1)	.03[b]
Han Chinese	3336 (91.9)	541 (89.1)		
Other	296 (8.1)	66 (10.9)		
Residence, n (%)			2.796 (1)	.09[b]
Urban	540 (14.9)	74 (12.2)		
Rural	3092 (85.1)	533 (87.8)		
Marital status, n (%)			40.658 (4)	<.001[b]
Currently married and living with a spouse	759 (20.9)	62 (10.2)		
Separated	51 (1.4)	8 (1.3)		
Divorced	22 (0.6)	3 (0.5)		
Widowed	2751 (75.7)	529 (87.1)		
Never married	49 (1.3)	5 (0.8)		
Exercise habits, n (%)			0.027 (1)	.87[b]
Yes	990 (27.3)	168 (27.7)		
No	2642 (72.7)	439 (72.3)		
Smoking status, n (%)			32.136 (2)	<.001[b]
Current smoker	799 (22)	78 (12.9)		
Former smoker	600 (16.5)	89 (14.7)		
Nonsmoker	2233 (61.5)	440 (72.5)		
Alcohol use, n (%)			21.517 (2)	<.001[b]
Current drinker	943 (26)	120 (19.8)		
Former drinker	441 (12.1)	52 (8.6)		
Nondrinker	2248 (61.9)	435 (71.7)		
Self-reported diabetes, n (%)			<0.001 (1)	>.99[b]
Yes	23 (0.6)	4 (0.7)		
No	3609 (99.4)	603 (99.3)		
Self-reported hypertension, n (%)			5.042 (1)	.02[b]
Yes	502 (13.8)	63 (10.4)		
No	3130 (86.2)	544 (89.6)		
Self-reported stroke or CVD[c], n (%)			1.849 (1)	.17[b]
Yes	133 (3.7)	15 (2.5)		
No	3499 (96.3)	592 (97.5)		
Self-reported cancer, n (%)			0.528 (1)	.47[b]
Yes	17 (0.5)	1 (0.2)		
No	3615 (99.5)	606 (99.8)		

Characteristics	Noncentenarians (n=3632)	Centenarians (n=607)	Statistic (*df*)	*P* value
Edentulousness, n (%)			23.847 (1)	<.001[b]
Yes	1206 (33.2)	264 (43.5)		
No	2426 (66.8)	343 (56.5)		
Denture use, n (%)			15.562 (1)	<.001[b]
Yes	784 (21.6)	88 (14.5)		
No	2848 (78.4)	519 (85.5)		

[a]Kruskal-Wallis H test.

[b]Chi-square test.

[c]CVD: cardiovascular disease.

Correlation Between Edentulousness and the Likelihood of Becoming a Centenarian

A significant association between edentulousness and the likelihood of becoming a centenarian was found among the participants. After adjusting for all covariates in model 3, the likelihood for becoming a centenarian increased in the nonedentulous group compared to the edentulous group (odds ratio [OR] 1.384, 95% CI 1.093 - 1.751; *P*=.007; Table 2).

Table . Association between edentulousness and the likelihood of becoming a centenarian.

Edentulousness	Participants (N=4239), n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
		OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Yes	1470 (34.7)	Ref[e]		Ref		Ref	
No	2769 (65.3)	0.646 (0.542-0.769)	<.001	1.311 (1.042-1.648)	.02	1.384 (1.093-1.751)	.007

[a]Model 1: unadjusted.

[b]Model 2: model 1 plus additional adjustment for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits.

[c]Model 3: model 2 plus additional adjustment for denture use and disease histories (including diabetes, hypertension, stroke or cardiovascular disease, and cancer).

[d]OR: odds ratio.

[e]Ref: reference.

Sensitivity Analyses

The robustness of the observed association was confirmed in the sensitivity analyses. PSM analysis was used to further evaluate the association between edentulousness and the likelihood of becoming a centenarian. Participants were divided into two groups for PSM analysis—the edentulous group and the nonedentulous group—and nearly all covariates were not significantly different after PSM (Multimedia Appendix 1). A significant increase in the likelihood of becoming a centenarian in the nonedentulous group was observed compared to the reference group (OR 1.272, 95% CI 1.037 - 1.561; *P*=.02; Table 3). Even after excluding participants who were followed up for less than 2 years, the association remained significant (OR 1.522, 95% CI 1.083 - 2.140; *P*=.02; Table 4). Regarding the association between the number of natural teeth as a continuous variable and the likelihood of becoming a centenarian, a notable positive association was found after adjusting for all covariates (OR 1.022, 95% CI 1.002 - 1.042; *P*=.03; Multimedia Appendix 2), which was comparable with the main results of this study.

Table . Association between edentulousness and the likelihood of becoming a centenarian after propensity score matching.

Edentulousness	Participants (n=2560), n (%)	OR[a] (95% CI)	*P* value
Yes	1280 (50)	Ref[b]	
No	1280 (50)	1.272 (1.037-1.561)	.02

[a]OR: odds ratio.

[b]Ref: reference.

Table . Association between edentulousness and the likelihood of becoming a centenarian in a 2-year lag analysis.

Edentulousness	Participants (n=2310), n (%)	Model 1[a]		Model 2[b]		Model 3[c]	
		OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Yes	410 (17.7)	Ref[e]		Ref		Ref	
No	1900 (82.3)	0.712 (0.572-0.885)	.002	1.414 (1.017-1.966)	.04	1.522 (1.083-2.140)	.02

[a]Model 1: unadjusted.

[b]Model 2: model 1 plus additional adjustment for the age in 1998, sex, ethnicity, marital status, smoking status, alcohol use, and exercise habits.

[c]Model 3: model 2 plus additional adjustment for denture use and disease histories (including diabetes, hypertension, stroke or cardiovascular disease, and cancer).

[d]OR: odds ratio.

[e]Ref: reference.

Discussion

Principal Findings

In this large-scale cohort study, the association between edentulousness and the likelihood of becoming a centenarian among populations aged 80 years or older was comprehensively analyzed. The likelihood of becoming a centenarian increased for the nonedentulous group compared to the edentulous group. Importantly, the correlation persisted after PSM analysis and was confirmed by the 2-year lag analysis. A significant positive association was found between the number of natural teeth and the likelihood of becoming a centenarian, aligning with the primary findings of this investigation.

The outcomes of this study generally support previous research, suggesting that retaining natural teeth is linked to reduced mortality rates among individuals aged 80 years or older. A previous study indicated that a majority of centenarians expressed contentment with their oral health [19]. A cross-sectional study using data from the New England Centenarian Study unveiled a lower prevalence of tooth loss among centenarians compared to their peers at the ages of 65-74 years, hinting at tooth loss as a potential indicator of decreased longevity. Notably, both centenarians and their descendants exhibited superior oral health compared to the control group [8]. Another cross-sectional study that included 1034 centenarians from the CLHLS dataset indicated that having ≤20 natural teeth was an independent risk factor for frailty among centenarians [3], underscoring the significance of dental health for this age group. A population-based survey from Finland disclosed that even a small number of missing teeth could signify a heightened risk of overall mortality, emphasizing the link between tooth loss and mortality [20]. These findings align with the results of this study, and this research further enhances existing knowledge by showcasing that possessing natural teeth correlates with an increased likelihood of reaching centenarian status through extensive follow-ups and several sensitivity analyses. The robustness of this study was reaffirmed by a series of sensitivity analyses, as the association persisted after PSM analysis and was confirmed by the 2-year lag analysis. Upon examining the relationship between the number of natural teeth and the probability of attaining centenarian status, a significant positive correlation emerged. This association remained robust even after adjusting for all covariates, aligning with the principal findings of this study. These results underscore the importance of managing oral diseases, suggesting that older adults with natural teeth have a greater chance of becoming centenarians.

In this study, the OR changed from 0.646 to 1.384 from model 1 to model 3 for the correlation between edentulousness and the likelihood of becoming a centenarian, suggesting the existence of potential confounding factors. By adjusting the variables that were included in the model, age in 1998 was found to be the variable causing the OR to invert. Given that the outcome of this study was centenarian status, the age of the participants at enrollment could profoundly influence the outcome. Older individuals at enrollment were more likely to reach the age of 100 years. Therefore, to further mitigate this influence, a 2-year lag analysis was conducted in the *Sensitivity Analyses* section, excluding older individuals whose follow-ups were less than 2 years.

The precise mechanism underlying the higher likelihood of reaching centenarian status among individuals with natural teeth remains incompletely understood. Previous studies indicated an association between the number of teeth and masticatory function [21], and reduced chewing ability was related to premature death [22]. Individuals with edentulousness may have an unbalanced food selection, consuming inadequate amounts of fruits and vegetables, and their nutritional status may dispose these individuals to more chronic diseases [7,23], as the dietary intake pattern influences the microbial compositions and systemic inflammation [24].

This study demonstrates notable strengths by using longitudinal data from a sizable, nationally representative cohort of older Chinese individuals, enabling a prospective assessment of the link between edentulousness and achieving centenarian status in older adults. Furthermore, the study used various models that adjusted for multiple variables and conducted several sensitivity analyses to enhance the reliability of the results. Significantly, this research unveils, for the first time, the influence of edentulousness on the likelihood of becoming a centenarian. Nevertheless, this study is subject to various limitations. First, most lifestyle behaviors were self-reported, introducing potential measurement errors. Second, medical conditions were also self-reported, despite detailed explanations provided by trained interviewers during data collection. This method likely led to

an underestimation of disease prevalence, possible misclassification, and residual confounding due to unmeasured medical conditions [6]. Third, despite adjusting for numerous confounding factors, the association could still be influenced by unmeasured or residual confounders, such as nutrition status that may affect both dental conditions and lifespan. Fourth, there may be reverse causality in this study. Individuals in better health status are more likely to retain their teeth, potentially biasing the interpretation of the findings. Finally, the survivorship bias may exist, as the participants included were aged ≥80 years at baseline, which may represent a population in relatively better health and may not be fully representative of all older adults.

Conclusions

In this cohort study involving individuals aged 80 years or older in China, the presence of natural teeth was linked to an increased probability of reaching the age of 100 years, emphasizing the significance of preserving oral health even in advanced age. Implementing targeted intervention strategies for oral health to enhance overall well-being could potentially contribute to longevity. Further prospective research and basic research experiments are essential to validate these findings and illuminate the underlying mechanisms.

Acknowledgments

This study was supported by the Shuguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (22SG14), the Shanghai Rising Star Project (22QA1405200), the Biomaterials Joint Research Project (2022LHB03), the Biobank Project of the Ninth People's Hospital–affiliated Shanghai Jiao Tong University School of Medicine (YBKA202207), and the National Natural Science Foundation of China (82201094).

Data Availability

The Chinese Longitudinal Healthy Longevity Survey data of this study are available on the web [25].

Authors' Contributions

HL and BL are co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the participants after propensity score matching (n=2560).
[DOCX File, 24 KB - aging_v8i1e68444_app1.docx]

Multimedia Appendix 2
Association between the number of natural teeth and the likelihood of becoming a centenarian.
[DOCX File, 17 KB - aging_v8i1e68444_app2.docx]

References

1. Yang J, Bai R, Wang L, Feng X, Liu Y, Zhou M. A projection of life expectancy based on the Global Burden of Disease Study 2019 - China, 1990-2030. China CDC Wkly 2023 Jan 13;5(2):40-44. [doi: 10.46234/ccdcw2023.008] [Medline: 36776687]
2. GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950-2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1160-1203. [doi: 10.1016/S0140-6736(20)30977-6] [Medline: 33069325]
3. Zhang J, Xu L. Frailty and associated factors among Chinese centenarians. J Nutr Health Aging 2022;26(8):806-813. [doi: 10.1007/s12603-022-1831-7] [Medline: 35934826]
4. Botelho J, Mascarenhas P, Viana J, et al. An umbrella review of the evidence linking oral health and systemic noncommunicable diseases. Nat Commun 2022 Dec 9;13(1):7614. [doi: 10.1038/s41467-022-35337-8] [Medline: 36494387]
5. Peres MA, Macpherson LMD, Weyant RJ, et al. Oral diseases: a global public health challenge. Lancet 2019 Jul 20;394(10194):249-260. [doi: 10.1016/S0140-6736(19)31146-8] [Medline: 31327369]
6. Li Y, Wang K, Jigeer G, et al. Healthy lifestyle and the likelihood of becoming a centenarian. JAMA Netw Open 2024 Jun 3;7(6):e2417931. [doi: 10.1001/jamanetworkopen.2024.17931] [Medline: 38900423]
7. Friedman PK, Lamster IB. Tooth loss as a predictor of shortened longevity: exploring the hypothesis. Periodontol 2000 2016 Oct;72(1):142-152. [doi: 10.1111/prd.12128] [Medline: 27501497]

8. Kaufman LB, Setiono TK, Doros G, et al. An oral health study of centenarians and children of centenarians. J Am Geriatr Soc 2014 Jun;62(6):1168-1173. [doi: 10.1111/jgs.12842] [Medline: 24889721]
9. Holm-Pedersen P, Schultz-Larsen K, Christiansen N, Avlund K. Tooth loss and subsequent disability and mortality in old age. J Am Geriatr Soc 2008 Mar;56(3):429-435. [doi: 10.1111/j.1532-5415.2007.01602.x] [Medline: 18194226]
10. Shimazaki Y, Soh I, Saito T, et al. Influence of dentition status on physical disability, mental impairment, and mortality in institutionalized elderly people. J Dent Res 2001 Jan;80(1):340-345. [doi: 10.1177/00220345010800010801] [Medline: 11269726]
11. Padilha DMP, Hilgert JB, Hugo FN, Bós AJG, Ferrucci L. Number of teeth and mortality risk in the Baltimore Longitudinal Study of Aging. J Gerontol A Biol Sci Med Sci 2008 Jul;63(7):739-744. [doi: 10.1093/gerona/63.7.739] [Medline: 18693229]
12. Borras C, Ingles M, Mas-Bargues C, et al. Centenarians: an excellent example of resilience for successful ageing. Mech Ageing Dev 2020 Mar;186:111199. [doi: 10.1016/j.mad.2019.111199] [Medline: 31899226]
13. Zeng Y, Shen K. Resilience significantly contributes to exceptional longevity. Curr Gerontol Geriatr Res 2010;2010:525693. [doi: 10.1155/2010/525693] [Medline: 21197075]
14. GBD 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1684-1735. [doi: 10.1016/S0140-6736(18)31891-9] [Medline: 30496102]
15. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr 22;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2] [Medline: 28285816]
16. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X]
17. Corrado OJ. Dentures. BMJ 1990 Dec 1;301(6763):1265-1268. [doi: 10.1136/bmj.301.6763.1265] [Medline: 2271830]
18. Ho D, Imai K, King G, Stuart EA. MatchIt: nonparametric preprocessing for parametric causal inference. J Stat Softw 2011 Jun 14;42(8):1-28. [doi: 10.18637/jss.v042.i08]
19. Rupel K, Tettamanti M, Vella F, et al. What do we learn from the clinical and biological evaluation of the oral cavity in centenarians? Maturitas 2021 Mar;145:31-37. [doi: 10.1016/j.maturitas.2020.12.005] [Medline: 33541560]
20. Liljestrand JM, Havulinna AS, Paju S, Männistö S, Salomaa V, Pussinen PJ. Missing teeth predict incident cardiovascular events, diabetes, and death. J Dent Res 2015 Aug;94(8):1055-1062. [doi: 10.1177/0022034515586352] [Medline: 25991651]
21. Ueno M, Yanagisawa T, Shinada K, Ohara S, Kawaguchi Y. Masticatory ability and functional tooth units in Japanese adults. J Oral Rehabil 2008 May;35(5):337-344. [doi: 10.1111/j.1365-2842.2008.01847.x] [Medline: 18405269]
22. Wei X, Zhang X, Chen R, et al. Diet-related diseases mediate the effect of masticatory function on premature death in older adults. J Periodontal Res 2024 Aug 13. [doi: 10.1111/jre.13335] [Medline: 39136162]
23. Smit MJ, Nijholt W, Bakker MH, Visser A. The predictive value of masticatory function for adverse health outcomes in older adults: a systematic review. J Nutr Health Aging 2024 May;28(5):100210. [doi: 10.1016/j.jnha.2024.100210] [Medline: 38489994]
24. Agurs-Collins T, Alvidrez J, ElShourbagy Ferreira S, et al. Perspective: nutrition health disparities framework: a model to advance health equity. Adv Nutr 2024 Apr;15(4):100194. [doi: 10.1016/j.advnut.2024.100194] [Medline: 38616067]
25. Peking University Open Research Data Platform. URL: https://opendata.pku.edu.cn/ [accessed 2025-03-14]

Abbreviations

CLHLS: Chinese Longitudinal Healthy Longevity Survey
CVD: cardiovascular disease
OR: odd ratio
PSM: propensity score matching
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by F Puga; submitted 06.11.24; peer-reviewed by C Liu, P Yin; revised version received 18.01.25; accepted 05.03.25; published 21.03.25.

Please cite as:
Wei X, Zhuang L, Li Y, Shi J, Yang Y, Lai H, Liu B
Edentulousness and the Likelihood of Becoming a Centenarian: Longitudinal Observational Study
JMIR Aging 2025;8:e68444
URL: https://aging.jmir.org/2025/1/e68444
doi:10.2196/68444

Artificial Intelligence-Driven Biological Age Prediction Model Using Comprehensive Health Checkup Data: Development and Validation Study

Chang-Uk Jeong[1*], BS; Jacob S Leiby[1*], PhD; Dokyoon Kim[1,2], PhD; Eun Kyung Choe[3,4], MD, PhD

[1]Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA

[3]Department of Surgery, Seoul National University Hospital Healthcare System Gangnam Center, 38-40FL Gangnam Finance Center(prior Star Tower) 152, Teheran-ro, Gangnam-gu, Seoul, Republic of Korea

[4]Department of Surgery, College of Medicine, Seoul National University, Seoul, South Korea

[*]these authors contributed equally

Corresponding Author:
Eun Kyung Choe, MD, PhD
Department of Surgery, Seoul National University Hospital Healthcare System Gangnam Center, 38-40FL Gangnam Finance Center(prior Star Tower) 152, Teheran-ro, Gangnam-gu, Seoul, Republic of Korea

Abstract

Background: The global increase in life expectancy has not shown a similar rise in healthy life expectancy. Accurate assessment of biological aging is crucial for mitigating diseases and socioeconomic burdens associated with aging. Current biological age prediction models are limited by their reliance on conventional statistical methods and constrained clinical information.

Objective: This study aimed to develop and validate an aging clock model using artificial intelligence, based on comprehensive health check-up data, to predict biological age and assess its clinical relevance.

Methods: We used data from Koreans who underwent health checkups at the Seoul National University Hospital Gangnam Center as well as from the Korean Genome and Epidemiology Study. Our model incorporated 27 clinical factors and employed machine learning algorithms, including linear regression, least absolute shrinkage and selection operator, ridge regression, elastic net, random forest, support vector machine, gradient boosting, and K-nearest neighbors. Model performance was evaluated using adjusted R^2 and the mean squared error (MSE) values. Shapley Additive exPlanation (SHAP) analysis was conducted to interpret the model's predictions.

Results: The Gradient Boosting model achieved the best performance with a mean (SE) MSE of 4.219 (0.14) and a mean (SE) R^2 of 0.967 (0.001). SHAP analysis identified significant predictors of biological age, including kidney function markers, gender, glycated hemoglobin level, liver function markers, and anthropometric measurements. After adjusting for the chronological age, the predicted biological age showed strong associations with multiple clinical factors, such as metabolic status, body compositions, fatty liver, smoking status, and pulmonary function.

Conclusions: Our aging clock model demonstrates a high predictive accuracy and clinical relevance, offering a valuable tool for personalized health monitoring and intervention. The model's applicability in routine health checkups could enhance health management and promote regular health evaluations.

(JMIR Aging 2025;8:e64473) doi:10.2196/64473

KEYWORDS

biological age; aging clock; mortality; artificial intelligence; machine learning; record; history; health checkup; clinical relevance; gerontology; geriatric; older; elderly; aging; prediction; predictive; life expectancy; AI

Introduction

Over the past several decades, global life expectancy has increased remarkably, rising from 66.8 years in 2000 to 73.4 years in 2019, according to the World Health Organization. However, healthy life expectancy has not kept pace, increasing only from 58.3 years to 63.7 years during the same period [1]. This demographic shift toward an aging population has led to increased health care dependency and associated social costs. The medical industry related to aging and the social costs thereof are continuously increasing [2]. Accurately assessing biological aging is a critical first step in mitigating age-related diseases and their socioeconomic impact.

Biological age refers to an estimation of an individual's physiological and functional status, reflecting the cumulative effects of genetic, environmental, and lifestyle factors on the aging process [3]. An individual with a biological age younger than their chronological age may have a lower risk of developing age-related diseases, while an older biological age could indicate a higher vulnerability to such conditions. This highlights the clinical significance of accurately estimating the biological age for personalized health interventions and monitoring. While numerous studies have explored the human lifespan [4,5], their evaluation is challenging due to the required long-term observations and limited clinical applicability.

Applying findings from academic studies to clinical practice remains challenging. Current biological age prediction models, primarily based on conventional statistical methods such as multivariate regression analysis, rely on limited clinical data, restricting their predictive power and insights into the aging process [5-8]. Recent advances have led to models using omics data [9], including DNA methylation [10], transcriptome [11], metabolome [12], and telomere data [9]. However, these models face implementation challenges in clinical settings due to their complexity and the difficulty in measuring omics markers. The model requires multiple molecular modalities and functional data to exhibit a superior performance [9]. In addition, it is challenging to quantify the dynamic effects of environmental, lifestyle, behavioral, and interventional factors on biological age.

In Asian countries, the health checkup industry has been growing substantially [13-15], with individuals regularly monitoring their health status. However, these checkups typically only indicate normal or abnormal conditions for individual tests, lacking comprehensive health status indicators. Providing biological age predictions through an aging clock could serve as a valuable tool for the health screening of patients, offering a comprehensive health status measure and encouraging regular checkup participation.

This study investigated the clinical relevance of artificial intelligence (AI)-predicted biological age in the Korean population using comprehensive health checkup data, examining its relationship with various clinical characteristics.

Methods

Participants and Datasets

The study investigated the healthy population participating in comprehensive health checkups at the Seoul National University Hospital Gangnam center, from 2003 to 2016. The initial baseline data were used, and the participants included a total of 81,211 Koreans, who comprised the Health and Prevention Enhancement (H-PEACE) cohort. The details of the H-PEACE cohort have been described previously [16]. To summarize, each participant completed a questionnaire on their past medical history and underwent anthropometric measurements and laboratory tests after at least 10 hours of fasting on the same day. We also used the data from the Korean Genome and Epidemiology Study (KoGES) from the Korean Center for Disease Control and Prevention as the replication set. From the KoGES data, we used the health examination cohort (KoGES HEXA data), which included past medical history, anthropometric measurement, and laboratory data [17]. Briefly, the KoGES HEXA cohort is a national health examinee registry, consisting of 173,357 urban Korean adults who underwent health checkup programs. We used clinical factors overlapping between factors from the H-PEACE cohort data and the KoGES HEXA data. The exclusion criteria to define the healthy adult super-control cohort were as follows: (1) participants diagnosed with diabetes, hypertension, or dyslipidemia; (2) participants drinking alcohol more than 14 g/week; (3) current or previous smokers; (4) those aged less than 30 years; and (5) those having a history of malignant disease. To check the clinical implication of the predicted biological age, we performed multiclinical feature association study in the gene-environmental interaction and phenotype (GENIE) study [16], which consisted of 123 clinical factors and gene datasets. We used the 116 clinical factors, excluding the 27 factors used to predict biological age, to determine the multiple associations with predicted biological age.

Development of a Biological Age Prediction Model

This study aimed to develop a widely applicable biological age model using only basic health screening parameters. The clinical features used to construct the biological age model were based on routine clinical measurements, including standard demographic features, blood test results, and anthropometric measurements, rather than the findings from expensive specialized examinations. The selection of clinical factors was constrained by the requirement that they be present in both the training dataset (H-PEACE cohort) and the replication dataset (KoGES HEXA cohort). Through this feature selection process, we used 27 clinical factors as inputs for predicting the biological age, which are gender, anthropometric measurements (height, weight, BMI, and waist circumference), metabolic status (levels of fasting glucose, glycated hemoglobin [HbA1c], uric acid, total cholesterol, triglyceride, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol), liver functions (albumin, total bilirubin, alkaline phosphatase, aspartate aminotransferase, alanine aminotransferase, and gamma-glutamyl transferase), complete blood cell counts (white blood cell count, red blood cell count, hemoglobin, hematocrit, and platelet count), calcium and renal function (levels of blood urea nitrogen and creatinine, and the glomerular filtration rate calculated using the chronic kidney disease epidemiology collaboration (CKD-EPI) equation [18]).

The true labels were defined as the chronological age of the super-control population, based on the assumption that chronological age aligns with biological age in physiologically standard individuals. This cohort was carefully selected to exclude individuals with pathological conditions, such as metabolic diseases or malignant diseases, as well as those exposed to environmental factors known to influence biological age, including smoking and alcohol consumption [19].

The baseline data of the H-PEACE super-control cohort was split into 80% as the training set and 20% as the testing set with stratification based on both age and sex. Specifically, age stratification was conducted by categorizing participants into

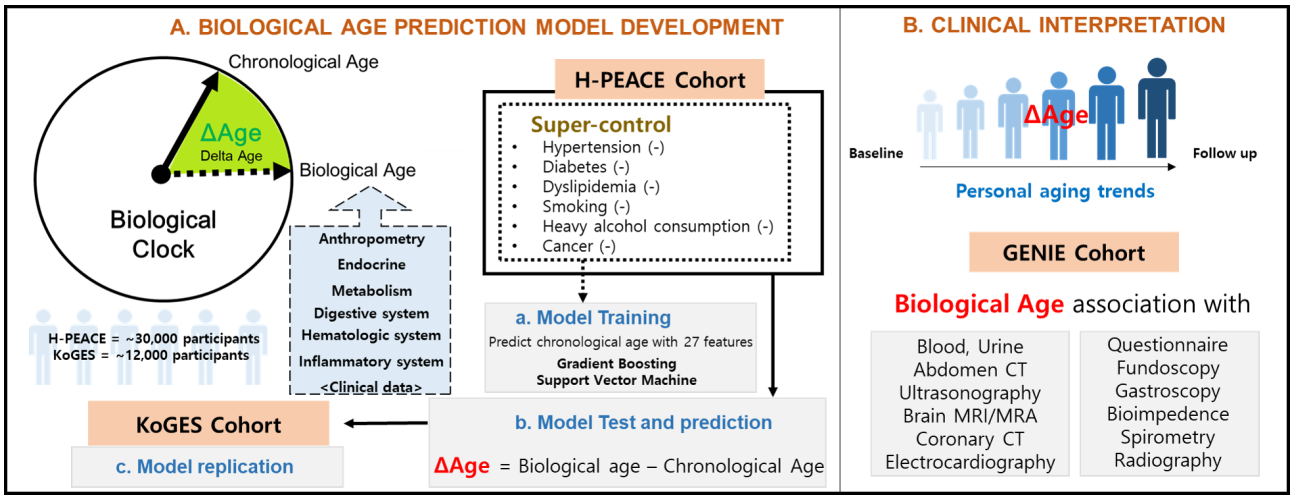
decade-based intervals, which were 20 - 29, 30 - 39, 40 - 49, 50 - 59, 60 - 69, 70 - 79, and more than 80 years. Five-fold cross-validation was performed to find the best hyperparameter for each model. We employed various machine learning models such as linear regression, least absolute shrinkage and selection operator (LASSO) regression, ridge regression, elastic net, random forest, support vector machine (SVM), gradient boosting, and K-nearest neighbors. These models were chosen for their ability to handle diverse relationships in data, including linear, nonlinear, and complex interactions. Hyperparameter optimization for each model was conducted using a grid search. To evaluate the performances of the developed models to predict biological age, we used adjusted R^2 and mean squared error (MSE) values. The evaluation results reported in this study are from 10 iterative experiments. A replication study was conducted in the KoGES HEXA dataset with the same experimental setting as the H-PEACE dataset. We interpreted the biological age prediction results using Shapley Additive exPlanation (SHAP) [20]. SHAP analysis was performed to elucidate the roles and impacts of different biological markers in predicting biological age.

Investigation of the Clinical Relevance of the Predicted Biological Age

We conducted linear regression analysis for multiple clinical factors using the GENIE study dataset with the predicted biological age, adjusting for chronological age. All reported *P* values were corrected for multiple tests using the Bonferroni correction.

The overview of the study is shown in Figure 1.

Figure 1. Overview of the study. H-PEACE: Health and Prevention Enhancement, KoGES: Korean Genome and Epidemiology Study, GENIE: gene-environmental interaction and phenotype, CT: computed tomography, MRI: magnetic resonance imaging, MRA: magnetic resonance angiography.



Ethical Considerations

The Institutional Review Board of the Seoul National University Hospital approved the study protocol and waived the need for informed consent (IRB number H-2005-223-1129). The study was performed in accordance with the Declaration of Helsinki. All patient data were deidentified before analysis, and strict confidentiality measures were implemented throughout the data collection, storage, and analysis processes. Access to the data was restricted to authorized research personnel only. The data management procedures complied with relevant data protection regulations to safeguard participants' privacy. No personally identifiable information was included in the final dataset or results.

Statistical and Computational Analyses

All analyses and calculations were performed using Python version 3.11.11 (Python Software Foundation). Multiple evaluation indices, including adjusted R^2 and MSE, β, and *P* value, were used to comprehensively evaluate the performances of the models and the significance of the associations. The statistical significance was based on a two-tailed *P* value of <.05.

Results

General Characteristics of the Participants

After applying the exclusion criteria to define the super-control cohort and removing the missing variables, we developed a model using a dataset collected from 28,417 individuals who underwent comprehensive health checkups at the Seoul National University Hospital Gangnam Center (ie, the H-PEACE cohort). The enrollment process is shown in Figure S1 in Multimedia Appendix 1. The mean (SD) participant age was 44.22 (11.26) years for 6467 men and 21,950 women. The baseline characteristics according to the aging of the study participants are shown in Table S1 in Multimedia Appendix 2. There were 1005 participants who were more than 65 years old and 27,412 who were 65 years old or less. Figure S2 in Multimedia Appendix 1 shows the chronological age distribution for the respective genders. Figure S3 in Multimedia Appendix 1 shows the distribution of gender in the training and test sets as well as the chronological age group distribution in the training and test sets.

Development and Performance of the Biological Age Prediction Model

The models were trained using 27 clinical variables on the full training dataset (80% of the super-control cohort), then predicted

on the test set (20%). For generalizability, 5-fold cross-validation with 10 iterative experiments were performed. The model was replicated using the KoGES HEXA data of 11,968 super-controls. Among 8 machine learning algorithms, the model showing the best performance was gradient boosting, for which the mean (SE) MSE value was 4.219 (0.140) and the mean (SE) R^2 value was 0.967 (0.001). The hyperparameters for the gradient boosting model were α=0.9, complexity parameter α=0, learning rate=0.1, maximum depth=5, and number of trees=500.

The second-best performing algorithm was the SVM model, with a mean (SE) MSE of 8.244 (0.210) and mean (SE) R^2 value of 0.935 (0.002). The performances of the 8 machine learning models in the test set are shown in Table 1.

Table . Comparison of the performances of 8 machine learning models to predict biological age in the test set (N=5684).

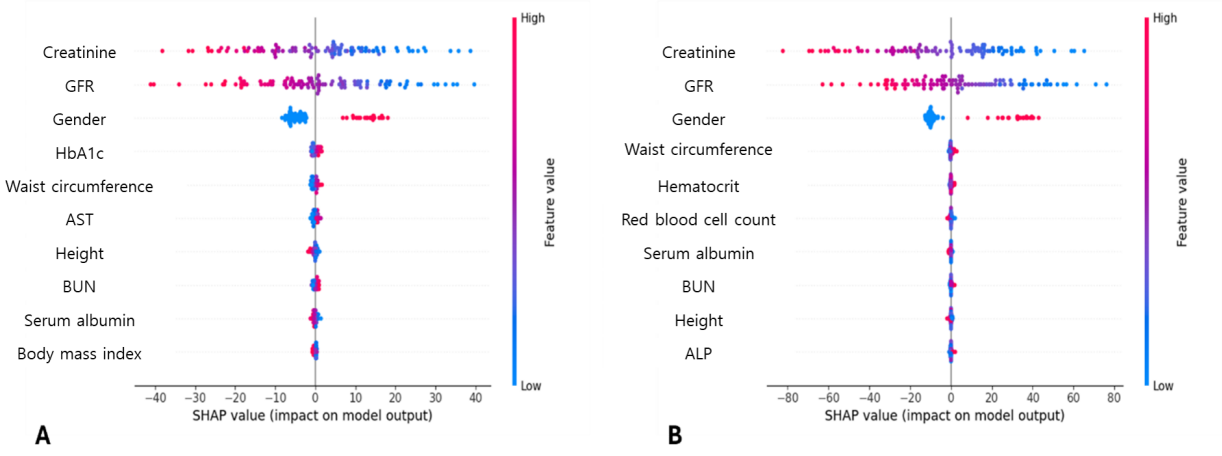
Model	Mean squared error, mean (standard error)	R^2, mean (standard error)
K nearest neighbor	63.829 (0.176)	0.497 (0.002)
Elastic net	50.518 (2.251)	0.602 (0.018)
Linear regression	50.314 (5.165)	0.603 (0.040)
LASSO[a]	50.271 (3.477)	0.604 (0.027)
Ridge	50.235 (5.047)	0.604 (0.039)
Random forest	20.941 (1.020)	0.835 (0.008)
Support vector machine	8.244 (0.210)	0.935 (0.002)
Gradient boosting	4.219 (0.140)	0.967 (0.001)

[a]LASSO: Least Absolute Shrinkage and Selection Operator.

The SHAP values from the gradient boosting and SVM models were analyzed to interpret their predictions of biological age. The SHAP values offer feature importance, providing the interpretation of the model's decision-making by quantifying the contribution of each feature to the model's output [20]. The corresponding visualizations are presented in Figure 2.

In the predictions of biological age generated by the gradient boosting model, the markers of kidney function, gender, HbA1c level, liver function, and anthropometric measurements were highlighted as significant predictors. In the SVM model, the SHAP summary plot revealed that kidney function markers, gender, liver function markers, red blood cell indices, and anthropometric measurements were the most influential predictors of biological age. These findings underscored the multifaceted nature of aging and highlighted the importance of maintaining optimal kidney function, metabolic status, and body composition in mitigating biological aging.

Figure 2. Visualization of the feature importance with Shapley Additive exPlanation (SHAP) values. SHAP summary plots for the model from the gradient boosting model (A) and the support vector machine (SVM) model (B) are visualized. Features with broader spreads and higher SHAP values have a more significant impact to predict biological age and values with the color gradient indicate whether higher or lower feature values are associated with increased biological age predictions. GFR: glomerular filtration rate, HbA1c: glycated hemoglobin, AST: aspartate aminotransferase, BUN: blood urea nitrogen, ALP: alkaline phosphatase.



Subgroup Analysis by Gender

The performances of the gradient boosting and SVM models were evaluated in the respective genders, male versus female participants. Applying the gradient boosting model to male participants, the mean (SE) MSE value was 5.258 (0.490). The SHAP value was significantly influenced by renal function, metabolic status, red blood cell indices, and anthropometric measurements. In female participants, the mean (SE) MSE value was 2.743 (0.099) and the SHAP values with the highest impact were similar to those for the male participants. The performances in the SVM and SHAP plots for male and female participants

are shown in Table 2 and Figure S4 in Multimedia Appendix 1.

Table . Comparison of the performances of the test set and replication set in all ages and in age above 65 years.

Dataset	Subgroup	Gradient boosting		Support Vector Machine	
		All ages	Age above 65 years	All ages	Age above 65 years
		MSE[a]	MSE[a]	MSE[a]	MSE[a]
Test set	Male and Female	4.219 (0.140)	18.942 (1.250)	8.244 (0.210)	18.077 (1.094)
Replication set	Male and Female	2.406 (0.083)	2.190 (0.085)	7.838 (0.065)	3.116 (0.148)
Test set	Male	5.258 (0.490)	20.357 (2.493)	15.476 (0.652)	20.678 (1.504)
Replication set	Male	2.032 (0.110)	2.096 (0.113)	6.337 (0.053)	3.143 (0.205)
Test set	Female	2.743 (0.099)	17.153 (0.819)	7.955 (0.170)	18.928 (1.698)
Replication set	Female	16.050 (0.467)	3.746 (0.366)	37.301 (0.777)	3.832 (0.510)

[a]Mean squared error (MSE) values are shown as mean (standard error).

Replication of the Developed Biological Age Prediction Model

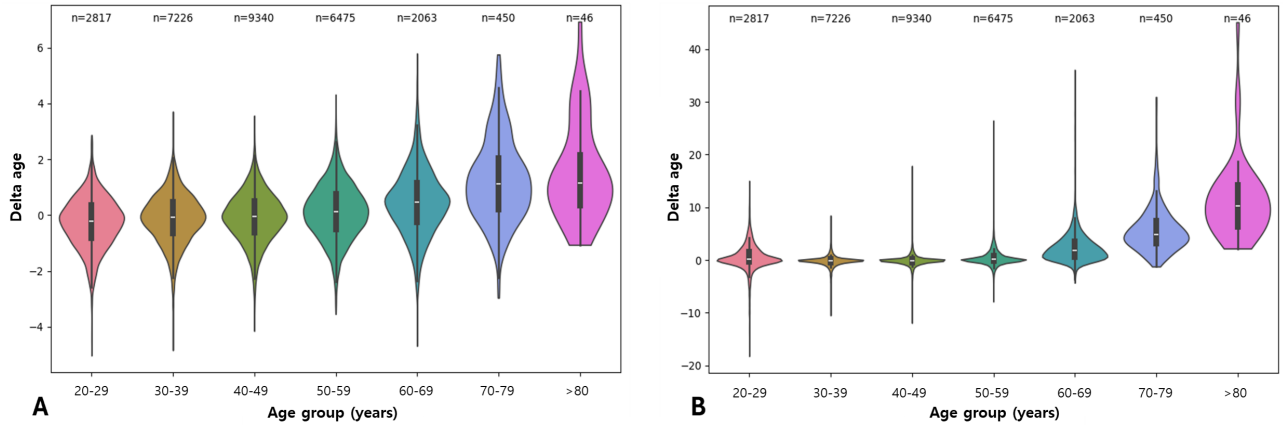
We replicated the model with the gradient boosting and SVM models using the KoGES HEXA data for participants of all ages and those under 65 years of age. The prediction model was better replicated in male participants and those aged above 65 years. The replicated results are shown in Table 2.

Delta Age Evaluation Using Baseline and Follow-Up Data

The model predicted the biological age in the test set (20% of the total dataset; 5684 individuals). The delta age, which is the difference between the biological age and the predicted biological age, was calculated for all individuals in the test dataset.

Figure 3 shows the delta age distribution across the different age groups, and Figure S5 in Multimedia Appendix 1 shows the delta age distribution across different age groups in each gender.

Figure 3. The delta age distribution across different age groups. The distribution of the delta age across different age groups in the test set at baseline (total n=5684) is shown for the gradient boosting (A) and support vector machine (B) models.



With the increase in the age groups, there was an observable pattern in the variability of delta age. In the younger age group (the 1st to the 3rd groups), there was increased variability, suggesting diverse aging processes influenced by multiple factors such as genetics, lifestyle, and health conditions. Older groups (the 7th and 8th groups) showed less variability and more positive values of delta age, possibly reflecting a selection of healthier individuals who have managed to reach older ages. The results were consistent in each gender group as well.

Among 5684 baseline participants, there were 2022 participants who had follow-up data, with follow-up duration ranging from 1 to 10 years (with 0 meaning the same age at two consecutive visits).

Figure 4 shows the projectiles of the changes in delta age during follow-up across different age groups in the test set (n=2022) for the gradient boosting and SVM models. While the younger age groups exhibited stable biological aging trajectories, the middle-aged and older groups showed increased variability and accelerated aging over the follow-up period. The observed trends were consistent across both the gradient boosting and SVM models, providing robust evidence for the described aging patterns.

Figure 4. The delta age trajectory during follow-up across different age groups. The projectiles of the changes in delta age during follow-up across different age groups in the test set (n=2022) are shown for the gradient boosting model (A) and SVM model.

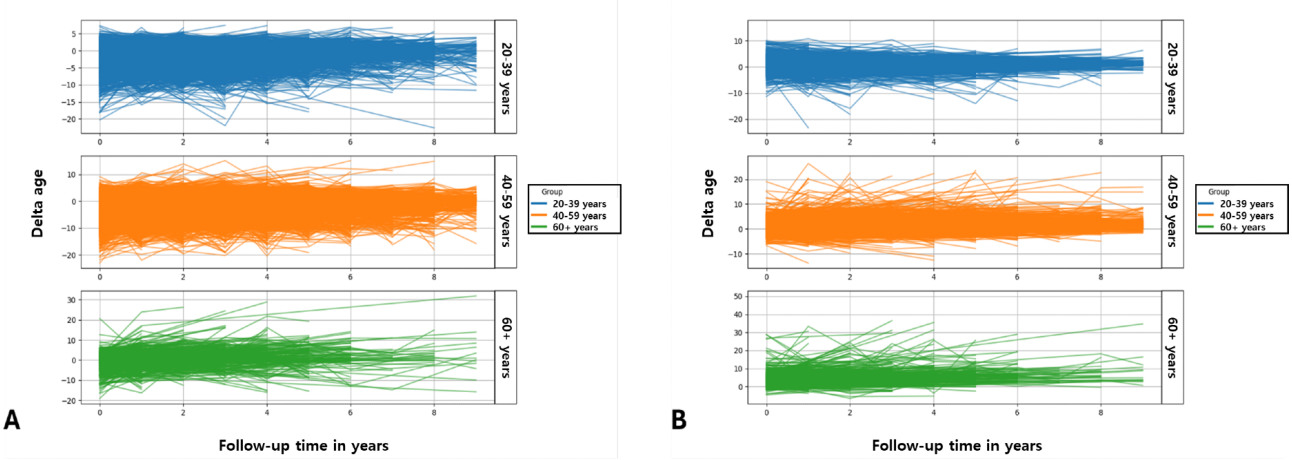


Figure S6 in Multimedia Appendix 1 shows the trajectory of the changes in delta age during follow-up across different gender groups in the test set at baseline (n= 5684) for the gradient boosting model and SVM model.

Clinical Interpretation Using a Multiple Clinical Factor Association Study

To explore the clinical relevance of the predicted biological age, we performed a phenome-wide association analysis between 116 phenotypes corroborated by comprehensive health checkups from the GENIE study dataset and the output of the predicted biological age model (data for both genders are shown in Table S2 in Multimedia Appendix 2, data for the male gender are shown in Table S3 in Multimedia Appendix 2, and data for the female gender are shown in Table S4 in Multimedia Appendix 2).

We found that the predicted biological age was significantly associated with 50 clinical factors after Bonferroni correction (Tables S2-S4 in Multimedia Appendix 2). Notably, the most significant associations were observed with metabolic status, body composition, fatty liver, smoking status, and pulmonary function, after adjusting for the chronological age. Table 3 shows the top 10 significant results in both genders, in male participants, and in female participants. The correlations among the variables are shown in Figure S7 in Multimedia Appendix 1. These findings suggest that the predicted biological age, developed from multiple clinical factors from comprehensive health checkups, may also possess predictive capabilities for aging in various organs.

Table . Top 10 significant clinical factors associated with the predicted biological age.

Clinical factors	Study modality	N	β	*P* value
In both genders				
Diabetes diagnosis	Questionnaire	10,351	−.002	1.75×10^{-270}
Skeletal muscle mass	Bioelectrical impedance	10,231	−.053	7.87×10^{-256}
Fatty liver	Ultrasonography	10,287	−.006	5.20×10^{-195}
Diabetes medication	Questionnaire	10,351	−.001	4.38×10^{-178}
Metabolic syndrome	Questionnaire	10,351	−.003	1.35×10^{-174}
Smoking status	Questionnaire	8995	−.004	8.43×10^{-170}
Visceral fat area	Abdominal computed tomography	6183	−46.952	2.86×10^{-153}
Forced vital capacity (liters)	Spirometry	10,138	−.005	6.32×10^{-141}
Forced expiratory volume in 1 s (liters)	Spirometry	10,138	−.004	5.71×10^{-135}
Chloride level	Blood	9964	.011	4.18×10^{-77}
In male participants				
Diabetes diagnosis	Questionnaire	4292	−.001	1.91×10^{-56}
Fatty liver	Ultrasonography	4270	−.004	2.74×10^{-46}
Metabolic syndrome	Questionnaire	4292	−.002	1.72×10^{-41}
Diabetes medication	Questionnaire	4292	−.001	9.63×10^{-36}
Mean corpuscular hemoglobin	Blood	4264	.013	3.61×10^{-35}
Mean corpuscular volume	Blood	4264	.026	1.54×10^{-27}
Visceral fat area	Abdominal computed tomography	2157	−26.868	3.61×10^{-24}
Total fat area	Abdominal computed tomography	2157	−48.823	1.23×10^{-14}
Mean corpuscular hemoglobin concentration	Blood	4264	.005	5.10×10^{-14}
Heart rate	Electrocardiography	3297	−.040	1.02×10^{-12}
In female participants				
Diabetes diagnosis	Questionnaire	6059	−.002	4.13×10^{-146}
Diabetes medication	Questionnaire	6059	−.002	2.49×10^{-98}
Metabolic syndrome	Questionnaire	6059	−.002	1.93×10^{-38}
Fatty liver	Ultrasonography	6017	−.004	5.03×10^{-31}
Potassium level	Blood	5839	.001	4.02×10^{-24}
Sodium level	Blood	5839	.007	1.59×10^{-23}
Chloride level	Blood	5839	.007	3.07×10^{-15}
Forced vital capacity (liters)	Spirometry	5956	.002	3.97×10^{-15}
Forced vital capacity percent	Spirometry	5956	.032	4.71×10^{-14}
Visceral fat area	Abdominal computed tomography	4026	−17.835	9.80×10^{-14}

Discussion

This study underscores the clinical relevance of biological age as predicted by AI using comprehensive health checkup data. Our findings elucidate the clinical relevance of biological age—as assessed through machine learning models—which exhibits strong associations with multiple clinical factors, thereby providing valuable insights into the aging process and its implications on health. The result demonstrated that the best-performing model, gradient boosting, achieved a mean (SE) MSE value of 4.219 (0.140) and mean (SE) R^2 value of 0.967 (0.001), indicating acceptable predictive accuracy. Additionally, the SHAP analysis highlighted markers such as kidney function, gender, HbA1c level, liver function, and anthropometric measurements as significant predictors of biological age. These findings suggest that biological age is a multifaceted construct influenced by various physiological factors and can serve as a robust indicator of the overall health status. The identification of these markers is significant because it supports the importance of maintaining optimal kidney function, metabolic health, and body composition in mitigating biological aging. This comprehensive understanding can lead to more targeted and effective interventions aimed at improving overall health and longevity.

Our study findings align with existing research findings that emphasize the utility of biological age as a comprehensive health indicator [21-23]. Previous studies have also identified the importance of factors such as kidney function [24], metabolic health [25,26], and inflammatory markers [27,28] in the aging process. However, our use of machine learning models such as the gradient boosting model and the SVM model provides a more nuanced understanding of how these factors interact to influence biological age, setting our research apart from traditional statistical methods. For instance, while existing literature has demonstrated the importance of individual biomarkers, our approach integrates multiple clinical variables to provide a holistic prediction model.

The use of multiple machine learning algorithms allowed for a comprehensive evaluation of their predictive capabilities. The gradient boosting model outperformed other algorithms, with performance R^2 of 0.967 and MSE of 4.219, due to its ability to handle nonlinear relationships and feature interactions, which are crucial in modeling complex biological systems [29]. Unlike linear models such as the ridge or LASSO regression, the gradient boosting model iteratively optimizes residual errors, capturing intricate dependencies between features. Its robustness to outliers and noise further enhances performance, making it particularly suited for real-world health checkups [30,31]. The generalizability of gradient boosting, demonstrated in both the training and replication datasets, underscores its potential for clinical application. Future studies should explore hybrid approaches to integrate the strengths of the gradient boosting model with more efficient models for broader use.

In the SHAP analysis, kidney function emerged as a significant predictor of biological age, consistent with its established role as an indicator of systemic health [32]. The decrease in kidney function is closely associated with aging-related changes, such as a reduced glomerular filtration rate and increased risk of chronic kidney disease [33]. These conditions often signify cumulative damage from metabolic stressors, hypertension, and other age-related factors [34]. The inclusion of kidney function markers, such as creatinine and the estimated glomerular filtration rate, in our model highlights their critical contribution to capturing the physiological aging process. Maintaining optimal kidney function could therefore serve as a target for mitigating biological aging and reducing the burden of associated comorbidities.

The body composition, including markers such as waist circumference and BMI, was another key contributor identified on SHAP analysis. This finding underscores the impact of adiposity and muscle mass on aging trajectories. Increased visceral fat and reduced skeletal muscle mass are hallmark features of sarcopenic obesity, a condition linked to metabolic dysfunction and accelerated biological aging [35,36]. Such changes exacerbate systemic inflammation and insulin resistance, further compounding aging-related risks [37,38]. The identification of body composition as a significant predictor emphasizes the importance of lifestyle interventions, such as exercise and dietary modifications, to preserve muscle mass and manage body fat. By integrating these markers into the prediction model, our study highlights their multifaceted roles in biological aging. Kidney function and body composition serve not only as indicators of systemic health but also as modifiable factors, offering potential avenues for personalized health interventions aimed at slowing biological aging and promoting longevity. This enhanced understanding enriches the clinical utility of our model and provides actionable insights for practitioners.

Visualization of the delta age distribution across different age groups shows that there is increased variability in the younger age group and less variability and a more positive value of delta age in the older group. This pattern underscores the importance of middle age as a critical period for implementing health interventions to manage biological aging effectively, and the need for focused care strategies to manage aging-related health issues in older adults.

In the multiple clinical factor association study for predicted biological age, we found that the predicted biological age was significantly associated with 50 clinical factors even after the adjustment of chronological age. The most significant associations were observed with the metabolic status, body composition, fatty liver, smoking status, and pulmonary function. These associations highlight the clinical relevance of biological age as a comprehensive marker of an individual's health status and suggest potential pathophysiological mechanisms accompanying the aging process. Reduced pulmonary function is known to be an indicator of systemic aging and has been linked to increased mortality and morbidity in older adults [39,40]. This suggests that maintaining optimal lung function could be crucial in mitigating the effects of biological aging and improving overall health outcomes.

The association between smoking status and predicted biological age is well-documented in various studies. For example, research using AI to analyze blood data from 149,000 adults revealed

that smokers exhibit a faster rate of biological aging compared to nonsmokers [41]. Remarkably, female smokers were predicted to have a biological age twice that of their chronological age [42]. These findings emphasize the detrimental impact of smoking on biological age. Additionally, exposure to tobacco during early life has been linked to accelerated biological aging in adulthood [43]. The underlying mechanisms include inflammation [44,45], epigenetic changes [46], and chromosomal damage [47]. These results underline the potential for biological recovery after cessation, aligning with prior research and emphasizing both the dangers of smoking and the benefits of quitting.

Body composition changes, including increased body fat, decreased muscle mass, and higher BMI, were also significantly associated with biological age. These findings align with previous studies that have shown that sarcopenia (loss of muscle mass) and obesity are critical factors in the aging process [48]. Excess body fat, particularly visceral fat, is associated with metabolic syndrome, cardiovascular diseases, and insulin resistance, all of which contribute to accelerated biological aging [49,50]. Conversely, maintaining muscle mass is essential for physical function and metabolic health, highlighting the importance of exercise and nutrition in managing the biological age. These findings also provide insights into the pathophysiological aspects of aging, suggesting that targeted interventions in these areas could potentially mitigate the adverse effects of aging and improve health outcomes.

The results have significant implications for health care and aging research. By providing a comprehensive measure of biological age, health care providers can better assess an individual's overall health status and identify potential risks for age-related diseases. This approach could also facilitate more personalized health care strategies, improving patient outcomes and quality of life. Furthermore, the findings contribute to the broader understanding of the aging process, offering valuable insights for researchers aiming to develop innovative strategies to overcome age-related diseases and enhance healthy aging.

Despite the robust findings, there are limitations to consider. The study population was predominantly Korean, which may limit the generalizability of the results to other ethnic groups. Additionally, the reliance on clinical data from health checkups may not capture all factors influencing biological age, such as genetic predispositions or epigenetic clocks and environmental influences. Future studies should aim to include more diverse populations and incorporate additional variables to enhance the comprehensiveness and applicability of the models. Furthermore, longitudinal studies are needed to validate the long-term predictive capabilities of the models and their effectiveness in different health care settings.

In conclusion, our study demonstrates the clinical relevance of biological age as predicted by machine learning models. The findings provide valuable insights into the aging process and highlight the potential of biological age as a comprehensive health indicator. Future research should focus on refining these models and exploring their applicability in diverse populations to further enhance their utility in clinical practice. The integration of advanced machine learning techniques with comprehensive health data holds great promise for advancing our understanding of aging and improving health care outcomes.

Acknowledgments

This study received funding from the Seoul National University Hospital Research Fund (grant 04-2020-2100).

Data Availability

Data are not publicly available due to restrictions (institutional policy to protect the privacy of research participants), but are available from the corresponding author on reasonable request.

Authors' Contributions

EKC and DK equally contributed as corresponding authors, and DK can be reached at dokyoon.kim@pennmedicine.upenn.edu for this purpose.
Data curation: EKC
Conceptualization and methodology: EKC and DK
Writing, analysis and interpretation of data, and software: EKC, CJ, and JSL
Funding acquisition: EKC
Review and editing and visualization: EKC, DK, CJ, and JSL
Data acquisition: EKC
All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[PDF File, 1208 KB - aging_v8i1e64473_app1.pdf]

Multimedia Appendix 2
Supplementary tables.
[XLSX File, 36 KB - aging_v8i1e64473_app2.xlsx]

References

1. World Health Organization. GHE: Life expectancy and healthy life expectancy. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-life-expectancy-and-healthy-life-expectancy [accessed 2024-06-20]
2. Chen J, Zhao M, Zhou R, Ou W, Yao P. How heavy is the medical expense burden among the older adults and what are the contributing factors? A literature review and problem-based analysis. Front Public Health 2023;11:1165381. [doi: 10.3389/fpubh.2023.1165381] [Medline: 37397714]
3. Wang Q, Wang Z, Mizuguchi K, Takao T. Biological age prediction using a novel DNN model based on steroid metabolic pathways. Bioinformatics. . [doi: 10.1101/2024.11.07.622443]
4. Sun Q, Yu D, Fan J, et al. Healthy lifestyle and life expectancy at age 30 years in the Chinese population: an observational study. Lancet Public Health 2022 Dec;7(12):e994-e1004. [doi: 10.1016/S2468-2667(22)00110-4] [Medline: 35926549]
5. Wang J, Chen C, Zhou J, et al. Healthy lifestyle in late-life, longevity genes, and life expectancy among older adults: a 20-year, population-based, prospective cohort study. Lancet Healthy Longev 2023 Oct;4(10):e535-e543. [doi: 10.1016/S2666-7568(23)00140-X] [Medline: 37804845]
6. Voitenko VP, Tokar AV. The assessment of biological age and sex differences of human aging. Exp Aging Res 1983;9(4):239-244. [doi: 10.1080/03610738308258458] [Medline: 6667707]
7. Takeda H, Inada H, Inoue M, Yoshikawa H, Abe H. Evaluation of biological age and physical age by multiple regression analysis. Med Inform (Lond) 1982;7(3):221-227. [doi: 10.3109/14639238209010720] [Medline: 7162237]
8. Krøll J, Saxtrup O. On the use of regression analysis for the estimation of human biological age. Biogerontology 2000;1(4):363-368. [doi: 10.1023/a:1026594602252] [Medline: 11708217]
9. Rutledge J, Oh H, Wyss-Coray T. Measuring biological age using omics data. Nat Rev Genet 2022 Dec;23(12):715-727. [doi: 10.1038/s41576-022-00511-7] [Medline: 35715611]
10. Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol 2019 Nov 25;20(1):249. [doi: 10.1186/s13059-019-1824-y] [Medline: 31767039]
11. Mao S, Su J, Wang L, Bo X, Li C, Chen H. A transcriptome-based single-cell biological age model and resource for tissue-specific aging measures. Genome Res 2023 Aug;33(8):1381-1394. [doi: 10.1101/gr.277491.122] [Medline: 37524436]
12. Lassen JK, Wang T, Nielsen KL, Hasselstrøm JB, Johannsen M, Villesen P. Large-scale metabolomics: predicting biological age using 10,133 routine untargeted LC-MS measurements. Aging Cell 2023 May;22(5):e13813. [doi: 10.1111/acel.13813] [Medline: 36935524]
13. Lu J. Ningen Dock: Japan's unique comprehensive health checkup system for early detection of disease. Glob Health Med 2022 Feb 28;4(1):9-13. [doi: 10.35772/ghm.2021.01109] [Medline: 35291200]
14. Gong W, Cheng KK. Challenges in screening and general health checks in China. Lancet Public Health 2022 Dec;7(12):e989-e990. [doi: 10.1016/S2468-2667(22)00207-9] [Medline: 36462520]
15. Kang HT. Current status of the national health screening programs in South Korea. Korean J Fam Med 2022 May;43(3):168-173. [doi: 10.4082/kjfm.22.0052] [Medline: 35610963]
16. Lee C, Choe EK, Choi JM, et al. Health and Prevention Enhancement (H-PEACE): a retrospective, population-based cohort study conducted at the Seoul National University Hospital Gangnam Center, Korea. BMJ Open 2018 Apr 19;8(4):e019327. [doi: 10.1136/bmjopen-2017-019327] [Medline: 29674364]
17. Kim Y, Han BG, GESg K. Cohort profile: the Korean Genome and Epidemiology Study (KoGES) consortium. Int J Epidemiol 2017 Apr 1;46(2):1350. [doi: 10.1093/ije/dyv316] [Medline: 27085081]
18. Fu EL, Levey AS, Coresh J, et al. Accuracy of GFR estimating equations based on creatinine, cystatin C or both in routine care. Nephrol Dial Transplant 2024 Mar 27;39(4):694-706. [doi: 10.1093/ndt/gfad219] [Medline: 37813817]
19. Husted KLS, Brink-Kjær A, Fogelstrøm M, et al. A model for estimating biological age from physiological biomarkers of healthy aging: cross-sectional study. JMIR Aging 2022 May 10;5(2):e35696. [doi: 10.2196/35696] [Medline: 35536617]
20. Wang K, Tian J, Zheng C, et al. Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP. Comput Biol Med 2021 Oct;137:104813. [doi: 10.1016/j.compbiomed.2021.104813] [Medline: 34481185]
21. Kang YG, Suh E, Lee JW, Kim DW, Cho KH, Bae CY. Biological age as a health index for mortality and major age-related disease incidence in Koreans: National Health Insurance Service - Health screening 11-year follow-up study. Clin Interv Aging 2018;13:429-436. [doi: 10.2147/CIA.S157014] [Medline: 29593385]
22. Jia Q, Chen C, Xu A, et al. A biological age model based on physical examination data to predict mortality in a Chinese population. iScience 2024 Mar 15;27(3):108891. [doi: 10.1016/j.isci.2024.108891] [Medline: 38384842]
23. Bafei SEC, Shen C. Biomarkers selection and mathematical modeling in biological age estimation. NPJ Aging 2023 Jul 1;9(1):13. [doi: 10.1038/s41514-023-00110-8] [Medline: 37393295]

24. Peng S, Xu R, Wei K, Liu N, Lv Y, Lin Y. Association between kidney function and biological age: a China health and retirement longitudinal study. Front Public Health 2023;11:1259074. [doi: 10.3389/fpubh.2023.1259074] [Medline: 38164447]
25. Bae CY, Piao M, Kim M, et al. Biological age and lifestyle in the diagnosis of metabolic syndrome: the NHIS health screening data, 2014-2015. Sci Rep 2021 Jan 11;11(1):444. [doi: 10.1038/s41598-020-79256-4] [Medline: 33431923]
26. Devrajani T, Abid S, Shaikh H, et al. Relationship between aging and control of metabolic syndrome with telomere shortening: a cross-sectional study. Sci Rep 2023 Oct 19;13(1):17878. [doi: 10.1038/s41598-023-44715-1] [Medline: 37857729]
27. Kozlov M. "Inflammation clock" can reveal body's biological age. Nature New Biol 2021 Jul 13. [doi: 10.1038/d41586-021-01915-x] [Medline: 34257449]
28. Li X, Li C, Zhang W, Wang Y, Qian P, Huang H. Inflammation and aging: signaling pathways and intervention therapies. Signal Transduct Target Ther 2023 Jun 8;8(1):239. [doi: 10.1038/s41392-023-01502-8] [Medline: 37291105]
29. Alghushairy O, Ali F, Alghamdi W, Khalid M, Alsini R, Asiry O. Machine learning-based model for accurate identification of druggable proteins using light extreme gradient boosting. J Biomol Struct Dyn 2024;42(22):12330-12341. [doi: 10.1080/07391102.2023.2269280] [Medline: 37850427]
30. Yarkın Yıldız A, Kalayci A. Gradient boosting decision trees on medical diagnosis over tabular data. arXiv. Preprint posted online on Oct 11, 2024. [doi: 10.48550/arXiv.2410.03705]
31. Singh R, Kumar A, Singh P, Sharma M. Improving cancer prediction in recommender systems through gradient boosting. Presented at: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT); Jun 24-26, 2024; Kamand, India p. 1-6. [doi: 10.1109/ICCCNT61001.2024.10724973]
32. Xu L, Yu C, Chen A, Li C, Mao Y. Longitudinal analysis of renal function changes in elderly populations: health status evaluation and risk factor assessment. Clin Interv Aging 2024;19:1217-1224. [doi: 10.2147/CIA.S450388] [Medline: 38974512]
33. Muglia L, Di Dio M, Filicetti E, et al. Biomarkers of chronic kidney disease in older individuals: navigating complexity in diagnosis. Front Med (Lausanne) 2024;11:1397160. [doi: 10.3389/fmed.2024.1397160] [Medline: 39055699]
34. Fossel M. Aging: How Aging Works, How We Reverse Aging, and Prospects for Curing Aging Diseases, 1st edition: Academic Press; 2024:1-17. [doi: 10.1016/B978-0-443-15500-0.00001-3]
35. Fermín-Martínez CA, Márquez-Salinas A, Guerra EC, et al. AnthropoAge, a novel approach to integrate body composition into the estimation of biological age. Aging Cell 2023 Jan;22(1):e13756. [doi: 10.1111/acel.13756] [Medline: 36547004]
36. Wei S, Nguyen TT, Zhang Y, Ryu D, Gariani K. Sarcopenic obesity: epidemiology, pathophysiology, cardiovascular disease, mortality, and management. Front Endocrinol (Lausanne) 2023;14:1185221. [doi: 10.3389/fendo.2023.1185221] [Medline: 37455897]
37. Aging Biomarker Consortium, Bao H, Cao J, et al. Biomarkers of aging. Sci China Life Sci 2023 May;66(5):893-1066. [doi: 10.1007/s11427-023-2305-0] [Medline: 37076725]
38. Briand M, Raffin J, Gonzalez-Bautista E, et al. Body composition and aging: cross-sectional results from the INSPIRE study in people 20 to 93 years old. Geroscience 2025 Feb;47(1):863-875. [doi: 10.1007/s11357-024-01245-6] [Medline: 39028455]
39. Sharma G, Goodwin J. Effect of aging on respiratory system physiology and immunology. Clin Interv Aging 2006;1(3):253-260. [doi: 10.2147/ciia.2006.1.3.253] [Medline: 18046878]
40. Thomas ET, Guppy M, Straus SE, Bell KJL, Glasziou P. Rate of normal lung function decline in ageing adults: a systematic review of prospective cohort studies. BMJ Open 2019 Jun 27;9(6):e028150. [doi: 10.1136/bmjopen-2018-028150] [Medline: 31248928]
41. Mamoshina P, Kochetov K, Cortese F, et al. Blood biochemistry analysis to detect smoking status and quantify accelerated aging in smokers. Sci Rep 2019 Jan 15;9(1):142. [doi: 10.1038/s41598-018-35704-w] [Medline: 30644411]
42. Skjodt N, Mamoshina P, Kochetov K, et al. Smoking causes early biological aging: a deep neural network analysis of common blood test results. 2018 Sep 15 Presented at: ERS International Congress 2018 abstracts; Sep 15, 2018. [doi: 10.1183/13993003.congress-2018.OA3809]
43. Cui F, Tang L, Li D, et al. Early-life exposure to tobacco, genetic susceptibility, and accelerated biological aging in adulthood. Sci Adv 2024 May 3;10(18):eadl3747. [doi: 10.1126/sciadv.adl3747] [Medline: 38701212]
44. Babizhayev MA, Yegorov YE. Smoking and health: association between telomere length and factors impacting on human disease, quality of life and life span in a large population-based cohort under the effect of smoking duration. Fundam Clin Pharmacol 2011 Aug;25(4):425-442. [doi: 10.1111/j.1472-8206.2010.00866.x] [Medline: 20698892]
45. Huang J, Okuka M, McLean M, Keefe DL, Liu L. Effects of cigarette smoke on fertilization and embryo development in vivo. Fertil Steril 2009 Oct;92(4):1456-1465. [doi: 10.1016/j.fertnstert.2008.07.1781] [Medline: 19019360]
46. Klopack ET, Carroll JE, Cole SW, Seeman TE, Crimmins EM. Lifetime exposure to smoking, epigenetic aging, and morbidity and mortality in older adults. Clin Epigenetics 2022 May 28;14(1):72. [doi: 10.1186/s13148-022-01286-8] [Medline: 35643537]
47. Mitiku T, Abebe B, Birhane N, Getie B. DNA Damage, Transposable Element Expression and Their Associated Factors in Aging: Preprints; 2024. [doi: 10.20944/preprints202410.1563.v1]

48. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
49. Bergman RN, Kim SP, Hsu IR, et al. Abdominal obesity: role in the pathophysiology of metabolic disease and cardiovascular risk. Am J Med 2007 Feb;120(2 Suppl 1):S3-S8. [doi: 10.1016/j.amjmed.2006.11.012] [Medline: 17296343]
50. Shah RV, Murthy VL, Abbasi SA, et al. Visceral adiposity and the risk of metabolic syndrome across body mass index: the MESA Study. JACC Cardiovasc Imaging 2014 Dec;7(12):1221-1235. [doi: 10.1016/j.jcmg.2014.07.017] [Medline: 25440591]

Abbreviations

AI: artificial intelligence
CKD-EPI: chronic kidney disease epidemiology collaboration
GENIE: gene-environmental interaction and phenotype
H-PEACE: Health and Prevention Enhancement
HbA1c: glycated hemoglobin
HEXA: health examination
KoGES: Korean Genome and Epidemiology Study
LASSO: least absolute shrinkage and selection operator
MSE: mean squared error
SHAP: Shapley Additive exPlanation
SVM: support vector machine

Edited by A Quialheiro; submitted 18.07.24; peer-reviewed by M Gasmi, M Malhotra, RS Gomaa Mahmoud, Y Cheng, Y Pan; revised version received 29.01.25; accepted 10.03.25; published 11.04.25.

Please cite as:
Jeong CU, Leiby JS, Kim D, Choe EK
Artificial Intelligence-Driven Biological Age Prediction Model Using Comprehensive Health Checkup Data: Development and Validation Study
JMIR Aging 2025;8:e64473
URL: https://aging.jmir.org/2025/1/e64473
doi:10.2196/64473

Peer Volunteers' Journeys Through Training and Engagement in Older Adult Communities: Descriptive Qualitative Study

Shefaly Shorey[1], PhD; Jeremy Soon Leong Seow[1]; Rathi Mahendran[2]; Geok Hua Wee[3]; Lee Gan Goh[4]; Ee Heok Kua[2]

[1]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Dr, Singapore, Singapore

[2]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[3]Principle Mindful Awareness Trainer (MAP), Age-Well Everyday Programme (AWE), Singapore, Singapore

[4]Division of Family Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Corresponding Author:
Shefaly Shorey, PhD
Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, 10 Medical Dr, Singapore, Singapore

Abstract

Background: The rising prevalence of mental health conditions such as depression and anxiety among the aging population underscores the need for accessible and effective psychosocial support, particularly for community-dwelling older adults who face barriers like social stigma and limited mental health literacy. Peer volunteers have emerged as a promising resource to support these individuals; yet, they often lack the requisite training for effective intervention.

Objective: This study aims to explore the experiences of peer volunteers who participated in a Psychological First Aid training program.

Methods: Using a descriptive qualitative research design, semistructured interviews were conducted with 13 older adults between September and October 2024, and data were thematically analyzed.

Results: Three themes were identified: (1) dimensions of volunteerism from motivations to resistance, (2) empowerment through collaborative learning, and (3) recommendations for designing inclusive, holistic training programs.

Conclusions: The findings of this study showed positive outcomes such as personal growth and strengthened social connections among participants. However, enhancements in teaching methods, logistical arrangements, and session regularity are recommended to optimize the Psychological First Aid program. These insights can guide the development of more robust training models to support both peer volunteers and the older adult communities they serve.

(JMIR Aging 2025;8:e71810) doi:10.2196/71810

KEYWORDS

psychological first aid; PFA training; peer volunteers; mental health; psychosocial support; older adults

Introduction

The global population of older adults has grown significantly in recent years. In Singapore, the proportion of citizens aged 65 years and older increased to 19.1% in 2023 and is projected to reach 24.1% by 2030 [1]. Globally, approximately 14% of older adults live with mental health conditions, with depression and anxiety being the most common [2]. This trend is mirrored in Singapore, where issues such as depression, dementia, and subsyndromal depression are increasingly prevalent among older adults [3]. Social isolation, multimorbidity, and low socioeconomic status have been identified as key contributors to the development of mental health conditions in older adults. These factors are associated with severe consequences, including increased suicide risk, diminished quality of life, and heightened susceptibility to cognitive disorders such as dementia [4-7]. Unfortunately, mental health conditions in this population often remain undiagnosed or underreported due to stigma and limited mental health literacy, leading to delays in treatment and intervention [8,9].

Volunteerism has been recognized as an effective approach to improving the mental well-being of older adults. Programs such as AmeriCorps Seniors and Age UK internationally, and SG Cares and Lions Befrienders locally, provide platforms for older adults to engage socially and meaningfully, helping to reduce social isolation and promote active aging [10-15]. Despite these benefits, studies conducted in Portugal and the United Kingdom reveal that volunteers often face challenges due to insufficient training to support individuals with mental health conditions [16,17]. Structured training programs have been increasingly recognized for their ability to prepare peer volunteers to provide effective psychosocial support. Peer volunteers' training has been widely used in various settings to help individuals cope with potentially traumatic events, reduce distress, and foster a

sense of safety while supporting others [18,19]. Such programs have also proven beneficial not only in humanitarian aid situations but also in enhancing volunteers' ability to offer timely, empathetic, and informal support [20,21]. Formal training equips volunteers with the knowledge and skills necessary to provide effective support to individuals in distress while also addressing the volunteers' own needs [22].

However, most studies examining peer volunteer training have been conducted in Western contexts, highlighting a critical gap in understanding its application and effectiveness in Asian settings [23-25]. Given that Asian cultures are deeply rooted in collectivism and family-centered caregiving, there is strong evidence that Western models shaped by individualism and structured volunteer systems may not fully translate to these contexts [26]. Considering the growing prevalence of mental health conditions among Singapore's older adults and the potential benefits of formal training programs, there is an urgent need to implement and evaluate such initiatives locally.

This study aims to explore the experiences of peer volunteers who participated in a Psychological First Aid (PFA) training program designed to support community-dwelling older adults in Singapore. Specifically, it addresses the research question: "What are the experiences of peer volunteers who completed the PFA programme?" The findings will inform the refinement of culturally relevant and effective training strategies to enhance psychosocial support for both volunteers and the older adult population in Singapore.

Methods

Study Design

A descriptive qualitative research design was used to explore the experiences of peer volunteers who participated in a PFA program. The use of a qualitative design is prudent as it allows participants to share "thick" or "rich" descriptions, providing deeper insights into subjective topics [27]. The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist guided the reporting of this qualitative study [28] (Table S1 in Multimedia Appendix 1).

Sampling Strategy and Eligibility

Convenience sampling was used to recruit eligible participants who met the following criteria: (1) older adults aged 50 to 95 years, (2) participated in the PFA program, and (3) spoke and read English.

PFA Program

The PFA program was developed based on foundational literature [29-31] and insights from a multidisciplinary research team. This team comprised psychiatrists, psychologists, nurses, and peer volunteers with extensive experience supporting community-dwelling older adults. Their prior primary and secondary research studies formed the basis of the PFA intervention design.

The program consisted of 4 half-day workshops, each lasting 4 hours, and conducted face-to-face over consecutive weekends. The workshops were held in a conducive community space and addressed a range of topics, including common mental health disorders in Singapore, psychosocial needs of older adults and peer volunteers, skills in empathetic listening, low-intensity psychological support, mindfulness training, identifying individuals at risk, and strategies for facilitating connections to professional care.

Workshops were delivered by a diverse team of trainers, including psychiatrists, nurses, experienced peer volunteers, and mindfulness instructors. Peer volunteers involved in earlier training sessions also contributed as facilitators, adding a practical perspective to the training content. The primary objective of the PFA program was to enable early detection of mental health issues among community-dwelling older adults. By equipping volunteers with these skills, the program also aimed to enhance the resilience and coping strategies of both the older adults and the volunteers themselves.

Ethical Considerations

The study received ethical approval from the institutional review board of the National University of Singapore (NUS-IRB-2024 - 506), which serves as the ethics board for the participating institution. Prior to recruitment, participants were provided with detailed information sheets outlining the study's objectives and their roles. Written informed consent was obtained, and voluntary participation was emphasized. To ensure anonymity and confidentiality, each participant's data were coded using unique serial numbers ranging from 1 to 13.

Data Collection

Data collection for the study was conducted between September and October 2024. All 20 peer volunteers who participated in the first batch of the PFA program were invited to join the study. Invitations were extended by a study administrator independent of the research team. Volunteers who expressed interest were introduced to the research team. A male research assistant (independent of the PFA program and participants) explained the study details to the interested participants, provided the participant information sheet, and obtained written informed consent. Subsequently, one-on-one interviews were arranged. Participants could choose between web-based or in-person interviews based on their preference. Before the interview, sociodemographic data such as age, gender, citizenship, ethnicity, religion, marital status, number of children, education level, employment status, and volunteering duration were collected.

A male research assistant (JS) with a BSc (Hons) in Nursing and no dependent relationship with the participants, developed the semistructured interview guide guided by literature review [32,33] and conducted the interviews to ensure consistency. Additionally, the same research assistant recorded field notes during each session. The guide included open-ended questions focusing on participants' experiences with the PFA program and their volunteerism journey (eg, "Would you like to share any personal experiences or stories that inspired you to join this PFA programme?" and "How was your experience with the PFA programme?"). All interviews were conducted via Zoom (Zoom Video Communications), and consent was obtained from participants for audio and video recording. Transcripts were transcribed verbatim from the recordings. To ensure rigor, the

research assistant (JS) received qualitative research training under the guidance of a PhD-trained researcher (SS) with extensive qualitative experience. A pilot interview was conducted with one peer volunteer to evaluate the interview guide's flow and relevance, aligning it with the research objectives; no modifications were needed, and the data from this pilot interview was included in the analysis as it contributed valuable insights to the study. A total of 13 interviews were conducted with no attrition, with durations ranging from 36 to 101 minutes (mean of 59.5 min). Data saturation was reached after the 11th interview, as no new concepts emerged [34]. Two additional interviews were conducted to confirm saturation. The interview guide is provided in Table S2 in Multimedia Appendix 1.

Data Analysis

The data were analyzed using Braun and Clarke's [35] 6-step framework for inductive thematic analysis. The process began with the verbatim transcripts being imported into Microsoft Word for organization and management. Two researchers (JS and SS) independently familiarized themselves with the data by reading and rereading the transcripts to gain a thorough understanding. Initial codes were generated by identifying meaningful units of information and relevant concepts within the data. These codes were systematically applied across all transcripts. Through an iterative process, the researchers grouped the codes into broader categories to develop initial themes and subthemes. These themes reflected the key patterns and narratives emerging from the data (Table S3 in Multimedia Appendix 1). The coding and thematic development were continuously refined through team discussions to ensure accuracy and depth. Any discrepancies in interpretation were resolved through consensus, ensuring a comprehensive and cohesive representation of the findings.

Rigor

To ensure the rigor of this study's findings, credibility, transferability, dependability, and confirmability were prioritized, as outlined by Cypress [36]. Credibility was maintained through prolonged engagement with participants during data collection and investigator triangulation during data analysis. This allowed for multiple perspectives to be considered and verified against the data. Transferability was established by providing thick, rich descriptions of participants' experiences with the PFA program, supported by direct quotes. This approach enables readers to evaluate the applicability of the findings to other contexts. Dependability was reinforced by creating an audit trail, which included comprehensive documentation of interview transcripts, coding, and the analytical process. This ensures that the study can be independently reviewed and verified. Confirmability was achieved using a reflexive journal throughout the research process. Reflexivity, as highlighted by Teh and Lek [37], is a critical aspect of qualitative research that minimizes the influence of unconscious researcher biases during data interpretation. By maintaining this journal, the researchers ensured their reflections and decisions were transparent, bolstering the trustworthiness of the study [38].

Results

Overview

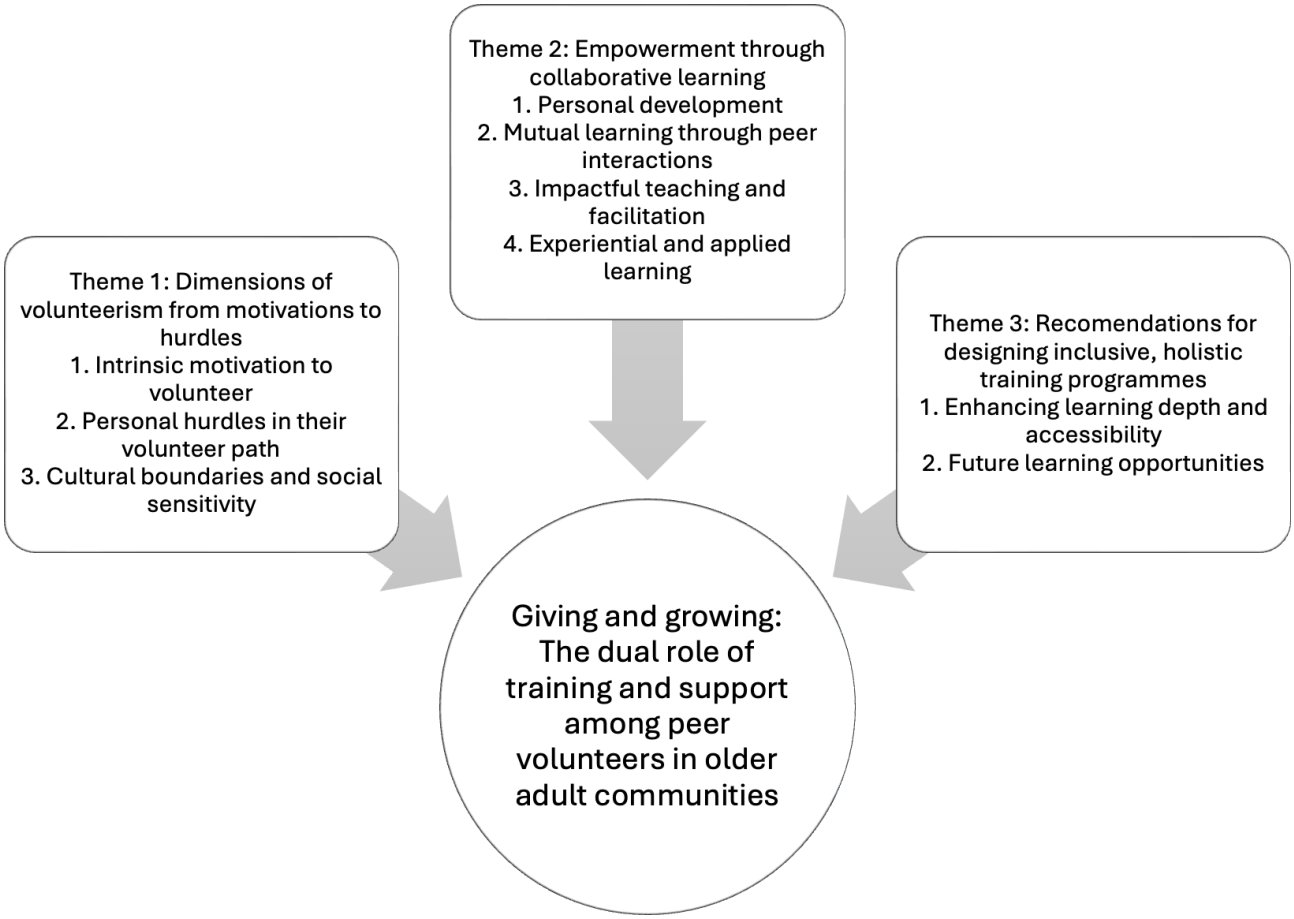
A total of 13 community-dwelling older adults participated in this study, and their sociodemographic characteristics are presented in Table 1. The mean age of the participants was 61.2 (SD 6.9) years. Most participants were female (9/13, 69%) and married (9/13, 69%), with all of them were Singaporeans (13/13, 100%). The group was predominantly composed of Malay (6/13, 46%) and Chinese (5/13, 38%) ethnic backgrounds. The participants' employment status varied, with being retired (4/13, 31%), employed full-time (4/13, 31%), employed part-time (2/13, 15%), unemployed (2/13, 15%), and self-employed (1/13, 8%). The group had a balanced distribution of volunteerism experience, with 54% (7/13) having up to 1 year of experience, and 46% (6/13) having more than 1 year of experience.

Three themes were identified from the data analysis, highlighting older adults' experiences of the PFA program and the value it added to their volunteerism journey: (1) dimensions of volunteerism from motivations to hurdles, (2) empowerment through collaborative learning, and (3) recommendations for designing inclusive, holistic training programs. These themes were further supported by 9 subthemes (Figure 1).

Table . Participants' sociodemographic characteristics (n=13).

Sociodemographic variables	Values
Age (years), mean (SD)	61.2 (6.9)
Sex, n (%)	
Female	9 (69)
Male	4 (31)
Nationality, n (%)	
Singaporean	13 (100)
Ethnicity, n (%)	
Chinese	5 (38)
Malay	6 (46)
Indian	1 (8)
Javanese	1 (8)
Marital status, n (%)	
Married	9 (69)
Single	3 (23)
Divorced	1 (8)
Length of marriage (years), n (%)	
0 - 10	4 (31)
11 - 20	1 (8)
21 - 30	2 (15)
31 - 40	3 (23)
>41	3 (23)
Highest education level, n (%)	
Secondary school	6 (46)
Polytechnic	2 (15)
University	5 (38)
Employment status, n (%)	
Employed full-time	4 (31)
Employed part-time	2 (15)
Self-employed	1 (8)
Unemployed	2 (15)
Retired	4 (31)
Volunteered as a peer volunteer (years), n (%)	
Up to 1	7 (54)
More than 1	6 (46)

Figure 1. Older adults' experience on the Psychological First Aid program and personal volunteering experience.



Theme 1: Dimensions of Volunteerism From Motivations to Hurdles

Overview

This theme highlighted older adults' experiences from motivations to hurdles, which shaped both their initial decision to volunteer and their continued commitment. These were categorized into 3 subthemes: intrinsic motivation to volunteer, personal hurdles in their volunteer path, and cultural boundaries and social sensitivity.

Intrinsic Motivation to Volunteer

Most older adults shared their personal motivations for volunteerism. As the majority of older adults were either employed part-time or retirees, they had free time available. Hence, they expressed a strong desire to actively engage with the community to enhance their personal well-being by giving back to society and fostering friendships. Many older adults also shared the perceived benefits of volunteerism, from keeping their minds active and serving as a protection to prevent the onset of dementia.

> *I volunteer as It's to help myself to prevent having dementia because I'm learning something new...* [K3, Female peer volunteer]

> *I volunteer as It's to help myself to prevent having dementia because I'm learning something new...* [K3, Female peer volunteer]

Some older adults attributed their volunteerism motivations to the cultural values instilled by their families. They were raised by their parents to be benevolent and gracious, which fostered a strong desire to give back to society. As a result, the concept of volunteerism became deeply embedded in their upbringing.

> *I see my father as a figure who do all this volunteer work in his life.… so it's already in my blood* [K5, Female peer volunteer]

Older adults attributed their decision to join the PFA program to the influence of their peers, highlighting the positive impact of fostering a sense of purpose and building meaningful friendships. For instance, one participant, who became a single mother during her teenage years, recounted the emotional struggles she faced, such as limited support from her family and peers, which negatively affected her emotional health. Having developed emotional resilience through overcoming these challenges, she felt inspired to volunteer and provide emotional support to others navigating similar difficulties.

> *To help people because I had gone through a period of time [her teenage years] and from that time, it makes me want to touch more and especially on single moms* [K13, Female peer volunteer]

Older adults shared that establishing meaningful engagement with their community residents was a significant challenge. Many encountered resistance when attempting to approach and offer support. Observing the reluctance of fellow older adults to seek help underscored for them the importance of acquiring

proper training. They recognized that such training was crucial to effectively support and connect with others in their communities.

> *I think it's the resistance of the people like when we try to talk or whatever, there is certain resistance… that led me to learn more from this programme* [K11, Female peer volunteer]

Personal Hurdles in Their Volunteer Path

Older adults highlighted the struggle of physical limitations, such as the lack of strength, despite being motivated to attend training courses to support others. This is an important consideration, as most community residents are older adults who may need assistance moving from one place to another. One participant shared that, due to her age and physical build, she is not as strong as she was in her youth and therefore needs to be more cautious when interacting with the older community residents or attending any courses.

> *I don't have much strength, things are very physical. I would not be able to help because … we may need to provide them physical support, like to support them so that they don't fall down...maybe emotional support would do better [for me]…* [K3, Female peer volunteer]

In addition, one participant shared that he was diagnosed with aphasia. This gravely impacted his communication abilities with both his peers and community residents. As a result, he struggled to express his needs efficiently.

> *Sometimes my train of thoughts go out... sometimes when people talk to me, I can't interpret properly… that sometime dampen me from opening up…* [K10, Male peer volunteer]

Cultural Boundaries and Social Sensitivity

Many older adults identified cultural perceptions as a significant barrier in their volunteerism experiences. They faced cultural taboos, such as Muslims' discomfort with certain terms used around older adults, including "forgetful" and concerns about interacting too much with the opposite gender. As a result, the volunteers, especially those of different ethnicities, had to be mindful of their behavior when approaching the community-dwelling older adults. Additionally, some volunteers highlighted that Muslim residents in the community often exhibit caution towards "outsiders" and show a preference for staying within their own social and ethnic circles, which limits their willingness to engage with others. These cultural dynamics present challenges in fostering broader social connections and offering support.

> *For the Malay elderly how we want to get them, how are we going to approach them to ask them that, we can really give them help. That's our challenge now.* [K13, Female peer volunteer]

Gender dynamics were frequently discussed by older adults, who reported that supporting members of the opposite gender was often challenging and required careful consideration, particularly regarding body language and posture. Many volunteers found it easier to assist individuals of the same gender, such as a female peer volunteer supporting a female older adult, as they could more easily interpret nonverbal cues like body language and gestures. One female peer volunteer shared her discomfort when a male community resident exhibited inappropriate behavior, such as touching without consent. These gender-related challenges highlight the complexities involved in volunteer interactions and the need for sensitive handling of such situations in community support roles.

> *Not easy to understand a male, what he's thinking. A female, when you look into her eyes, you talk to her, you can understand more… not easy with males….* [K1, Female peer volunteer]

Many felt the need for a chaperone, especially when interacting with the opposite gender, as a safety measure and to prevent unnecessary conflict. In one instance, one female peer volunteer experienced negative experiences from a wife due to her close rapport with the community resident.

> *There is a gender difference [in our local setting] that we need to be more mindful. For female, we can touch the female, can consult but for male, we need to be more careful, more mindful on our gesture.* [K11, Female peer volunteer]

Theme 2: Empowerment Through Collaborative Learning

Overview

This theme highlighted elements that contributed to the empowerment of older adults through collaborative learning during the PFA program. These were supported by 4 subthemes as follows: personal development, mutual learning through peer interactions, impactful teaching and facilitation, and experiential and applied learning.

Personal Development

Older adults highlighted the benefits of participating in the PFA program. Many expressed how the knowledge acquired through the PFA program helped them in real-life situations, enabling them to better showcase their skills as peer volunteers. Topics such as grief and mindfulness were essential in their role as peer volunteers. This, in turn, fostered a sense of empowerment and personal growth among the older adults.

> *It Was Nice, It Was Very Knowledgeable. I Must Say We Learn a Lot* [K1, Female Peer Volunteer]

Additionally, the older adults reported maintaining a positive attitude toward learning. They expressed that learning is a continuous journey, and the knowledge they gained expanded their viewpoints.

> *The… workshop where it really broadened our mind and then we learn more about things [eg how to support others better].* [K2, Male peer volunteer]

Mutual Learning Through Peer Interactions

Older adults emphasized that the PFA program highlighted the importance of maintaining open and effective communication among peers. Through interactions during class and breaks,

they had the opportunity to share personal experiences, which many participants found invaluable. This peer exchange allowed them to learn from each other's experiences and mentally prepare for similar real-life situations in the future. The older adults also expressed that the learning process was reciprocal, with everyone eager to contribute and learn from one another.

> *You get different ideas from them [peers attending PFA programme]. You learn more things from them, they learn from you.* [K1, Male peer volunteer]

Additionally, some older adults shared that the class, which included individuals from various ethnic backgrounds such as Chinese, Malay, and Indian, provided a valuable opportunity to engage with and learn about different cultural practices. This allowed them to conceptualize and compare the differences or similarities between the cultures.

> *The other group are Malays so then cultural wise, it's quite different. I think to understand each other's cultural difference… that was helpful* [K3, Female peer volunteer]

Impactful Teaching and Facilitation

Dedicated facilitators consistently worked to build rapport with the older adults, helping to ease initial discomfort and encouraging open communication. Older adults highlighted that the facilitators were highly knowledgeable in their respective fields. They were committed to sharing their knowledge on various topics such as handling difficult situations and mindfulness, which made the older adults feel comfortable asking questions whenever they had any doubts. With that, this fostered communication and the older adults were able to maximize the knowledge gained throughout this PFA program.

> *Think the way they [facilitators] present and they deliver the topic well, yeah and they are very knowledgeable on the topic… you feel motivated.* [K11, Female peer volunteer]

Additionally, many older adults highlighted that they were inspired by the facilitators' presentation style. The facilitators' consistent engagement, well-paced sessions, body language, tone, and delivery made the learning process more accessible and understandable for older learners. The older adults also valued the facilitators' sharing of personal experiences, as it allowed them to learn from real-life situations.

> *I learn from it and [blinded] because she's a mindfulness teacher and I'm learning to teach mindfulness. So of course I also want to learn all the right things from her.* [K3, Female peer volunteer]

Experiential and Applied Learning

Many older adults favored experiential learning over traditional lecture-style delivery. During the PFA program, they were able to actively participate in hands-on activities such as role-plays and this helped them to better visualize and retain the concepts taught.

> *The role play, it showed you like reality what could happen or what will happen or what can happen. That was really like an experience, very good because learning from slides and all is quite normal.* [K1, Female peer volunteer]

Additionally, participants reported that the role-playing exercises played a crucial role in boosting their confidence in handling real-life situations. By simulating real-world scenarios, they were able to "experience" these situations in a controlled environment. This allowed them to observe nonverbal cues and body language, which heightened their awareness of their own nonverbal communication when offering support. The interactive nature of the role-playing activities also facilitated communication among the older adults, helping to break the ice and foster connections.

> *When you have a role play, you really understand it more. You sometimes put yourself into that shoes and see how it is like.* [K1, Female peer volunteer]

Theme 3: Recommendations for Designing Inclusive, Holistic Training Programs

Overview

This theme highlighted recommendations to enhance the PFA program by addressing the varied needs, backgrounds, abilities, and learning styles of the participants. This theme is supported by 2 subthemes as follows: depth and quality of learning, and future learning opportunities.

Depth and Quality of Learning

A few older adults expressed that the content covered was too brief. They desired more in-depth knowledge to be taught, enabling them to gain additional insights that could be applied throughout their volunteerism journey. However, they noted that the brief content might be a result of the short session duration. Therefore, they expressed a preference for longer sessions.

> *I think it can go a bit deeper. It's just that the depth can be, in my personal opinion, if time allowed, the depth can be more* [K9, Female peer volunteer]

Furthermore, the older adults highlighted that the program venue was inconvenient. Some had to get up as early as 7 in the morning to arrive on time for the session. Despite the travel challenges, they expressed a desire for longer sessions to facilitate deeper learning.

> *I find it's really short because we [travel] from the east [and] we spend only two hours… the timing is not enough to share with us.* [K04, Female Peer Volunteer]

Many older adults envisage more hands-on practice during the PFA program. Though older adults caught up during breaks and role-plays, having a specific time for small-group discussions regularly was highlighted by some older adults.

> *I was thinking maybe it's good to have maybe small group discussion during each session.* [K11, Female peer volunteer]

A few older adults suggested incorporating various learning tools, such as case studies and better audio-visual materials for learning. Those older adults with less volunteer experience felt

the need for more case studies to be discussed so that they are more prepared to handle the situations in real life.

> *The other techniques like case study so examples of how we can approach the people… I never counsel anyone so I wouldn't know, what [is] the real situation outside... I am bit worried to handle if someone suddenly cry…* [K3, Female peer volunteer]

In addition, older adults highlighted several considerations that influenced their learning experience when engaging with older learners. They encountered issues with audio and visual clarity, such as some participants finding the projected fonts too small, requiring them to squint to read. Furthermore, the facilitators' lack of use of a microphone led to poor hearing, making it difficult for a few participants to fully comprehend the content taught.

> *As we age, our eyesight is not so good, we have to strain our eyes. I had to strain my eyes to read [the slides].* [K1, Female peer volunteer]

> *If you're seated behind far away, you can't hear much. I'm trying to figure out what he's saying. So if you have a mic, then at least, you know, we can hear, everybody can hear.* [K4, Female peer volunteer]

Future Learning Opportunities

Older adults highlighted the potential for future learning opportunities. Many expressed a desire for regular "top-up" sessions, either annually or semiannually, to revise the knowledge. Furthermore, they showed interest in gaining new knowledge through these regular sessions. A few older adults stressed the importance of regular sessions so that more peer volunteers from their respective organizations could participate and benefit from the PFA program.

> *It will be good if there's every six months or a year, we come back do skill upgrading.* [K9, Male peer volunteer]

> *I think we need to have more of this courses, for more people to be trained because now Singapore is aging.* [K7, Male peer volunteer]

In addition, they valued the regular sessions as an opportunity to reconnect with their peers. They were eager to meet and share their recent volunteerism experiences, fostering mutual learning through casual conversations.

> *A small group, then we will have regular meetups for sharing and to also impart certain message.* [K11, Female peer volunteer]

Discussion

Principal Findings

This study explored the experiences of peer volunteers in the PFA program in Singapore, providing insights into their motivations, challenges, and the overall impact of their participation. All interviewed peer volunteers participated in all 4 sessions of the program. The sample was primarily composed of Malay participants (n=6), followed by Chinese participants (n=5), with one participant from each of the Indian and Javanese communities. While Singapore's population is predominantly Chinese, Malay, and Indian, future studies could aim for a more balanced racial distribution to better reflect broader societal trends [39].

The older adults in this study cited various reasons for their continued involvement in volunteerism, including cultural values instilled by their families, peer influence, and the promotion of healthy aging. Through volunteerism, the older adults in this study were also able to actively engage with the community, fostering new friendships and a sense of purpose. This aligns with existing literature, with studies conducted in Japan, Australia, and the United Kingdom suggesting that the motivations for volunteerism in older adulthood were driven by community engagement, building new friendships, positive influences on personal well-being, and the desire to foster a sense of purpose [40-42]. Interestingly, altruistic attitudes, such as cultural values of graciousness instilled in them, were a key motivation for the older adults in our study. This contrasts with existing literature, where volunteers report motivations driven by egoistic reasons such as the development of a new skill and personal satisfaction [43,44]. This difference may be attributed to the strong Asian values that emphasize family cohesion and collective well-being [45]. As such, future studies can compare and explore volunteer motivations across different cultural contexts.

While the older adults in this study demonstrated a strong commitment to volunteerism, they also faced personal barriers, including difficulties in engaging with the community and physical limitations such as ongoing medical conditions or frailty. Chronic illnesses or frailty restrict their ability to participate in physically demanding tasks, limiting their ability to contribute as fully as they may have wished. This is consistent with studies conducted on older adults in Spain and Norway, which found that maintaining good health was a key factor enabling volunteerism as it was highly prioritized by older adults [46,47]. These challenges highlight the need for support systems that prioritize older adults' health to ensure their ongoing participation in volunteer activities, ultimately improving both their physical well-being and their ability to contribute meaningfully to the community.

Cultural and social sensitivity emerged as another key theme in this study. Older adults encountered cultural taboos, particularly within the Muslim community, where there was a preference for interactions within one's own ethnic or social circle. This limited broader community engagement, highlighting the importance of cultural sensitivity in volunteer programs. Additionally, gender dynamics were cited as a challenge, with female participants feeling more comfortable providing support to those of the same gender. These findings are consistent with research indicating that cultural and gender factors can influence interpersonal dynamics in volunteer settings [48]. This may be attributed to Singapore's diverse ethnic composition, where cultural norms are critical. Furthermore, in the Asian context, where gender dynamics are often less openly discussed, it is possible that the older adults in this study feel more at ease interacting with individuals of the same gender [49]. This suggests the importance of addressing these sensitivities in designing inclusive volunteer

programs that accommodate diverse cultural and gender preferences. Future research could explore how these factors shape volunteer experiences and inform the development of more culturally inclusive training programs.

Personal development emerged as another significant theme, with participants noting the value of peer interactions, role-plays, and experiential learning during the PFA program. These activities not only reinforced the content but also helped participants apply the concepts to real-world situations. This contrasts with a study of clinicians in the United States, where participants expressed a preference against interactive training methods [50]. The difference may be attributed to the older adult population's preference for active, hands-on learning approaches, which has been noted in other studies [51]. Facilitators also played a crucial role, providing reassurance and fostering a supportive learning environment. This finding is consistent with research indicating that facilitators' knowledge and support are integral to positive learning experiences [52,53]. Participants in this study also expressed a desire for more knowledge and confidence to handle real-life situations, aligning with findings from Japan and Haiti, where peer volunteer training boosted confidence in applying training strategies safely [54,55]. This highlights the importance of designing volunteer training programs that incorporate both theoretical knowledge and practical, experiential learning.

Designing inclusive training programs involves a holistic approach that considers accessibility, logistical support, practical learning experiences, and efficient course delivery, all of which contribute to enriched learning experiences and future improvement opportunities. Our study highlighted that the participants in our study desired greater comprehensive coverage of the content, with additional hands-on practice and small group discussions to enhance the training experience and add greater value to their volunteerism journey. Additionally, they expressed interest in incorporating other educational tools, such as case studies, to better assimilate real-world scenarios and mentally prepare themselves. This is consistent with studies conducted in Gaza, Korea, and Singapore, which emphasized that PFA trainers desire more in-depth training and new delivery approaches to psychologically equip themselves, enabling them to better support the recipients [25,52,56]. Additionally, older adults in our study emphasized the importance of regular sessions to stay informed about current information about PFA. Moreover, our study found that holding regular sessions would facilitate enrollment for prospective participants, providing them with comprehensive knowledge. These findings align with studies conducted in the United States and the United Kingdom, where participants envisage refresher sessions for PFA to further equip themselves with the latest knowledge [32,50]. Furthermore, several participants in our study, drawing on their prior volunteerism experience, expressed a desire for advanced volunteerism courses offering enhanced learning opportunities. This aligns with findings from studies conducted in the United States and Singapore, which highlight that individuals with volunteerism experience gain greater benefits from comprehensive training programs compared with basic ones [50,52]. Thus, a holistic approach to program design, which considers accessibility, practical learning opportunities, and continuous support, can enhance the overall success and impact of volunteer programs, and stakeholders can consider these recommendations in providing better support to the peer volunteers and, in turn, to the community-dwelling older adults.

Limitations

This study has a few limitations. First, this study was conducted on English-speaking older adults and may not be transferable to non–English-speaking older adults in Singapore. Furthermore, Indians were underrepresented in this study. Hence, future studies can consider exploring the purposive sampling technique to study participants who are non–English-speaking and from diverse ethnic backgrounds to get a holistic understanding of the phenomenon of interest. Second, interviews were conducted with a convenience sample of older adults who volunteered to participate and hence could have harbored particularly positive views from those participants who were motivated to share about the PFA program. Third, a potential limitation is that participants were not explicitly asked to provide feedback on the interview guide during the pilot interviews, specifically on aspects such as question clarity, sequencing, and overall structure. Such feedback could have offered valuable insights for refining the guide and enhancing its suitability. This should be considered in future research.

Implications for Future Research and Practice

Future research should explore the influence of cultural norms and gender dynamics in volunteer programs, particularly in diverse and multiethnic contexts like Singapore. Studies can investigate how cultural and gender considerations affect volunteer engagement and explore strategies for overcoming these barriers. Future studies should focus on the role of cultural sensitivity training and its impact on improving interactions among volunteers and community members. The physical limitations of older adults, including chronic illnesses or frailty, were identified as significant barriers to volunteerism. Future research should delve deeper into how health status influences volunteer participation and explore ways to accommodate volunteers with varying levels of physical ability. Moreover, longitudinal studies could assess the long-term health benefits of volunteerism and how these programs might mitigate the physical limitations of older adults. Stakeholders working with older adults, such as community partners and health care providers, could benefit from considering the cultural and age-related concerns highlighted in this study. Understanding the unique cultural values, gender dynamics, and health challenges faced by older adults is crucial for fostering more effective engagement and providing tailored support. This study highlighted the importance of hands-on learning, role-playing, and peer interactions in enhancing volunteers' understanding and confidence. Future research should focus on evaluating the effectiveness of different interactive learning methods across various populations. Additionally, exploring the impact of peer-led sessions versus professional-led training could help refine approaches to fostering confidence and competency in volunteers. Research into the most effective formats for delivering such programs—whether through in-person, via web, or hybrid methods—could contribute valuable knowledge to improve volunteer retention and engagement. This study

emphasizes that individuals with prior volunteerism experience benefit from a more comprehensive training curriculum. Hence, future research can explore the potential of offering both basic and advanced PFA programs for learners at different proficiency levels to enhance individual learning outcomes. While this study focused on short-term engagement with the PFA program, future research should explore the long-term effects of participation on older adults' well-being, volunteer retention, and community impact.

Conclusions

This study explored the experiences of older adult volunteers who participated in a PFA program in Singapore. Findings highlighted the complex interplay of cultural, social, and personal factors that influence the experiences of older adult volunteers. The findings suggest that training programs could be designed with cultural sensitivity, provide experiential learning opportunities, and support participants' ongoing personal development to maximize their engagement and impact. Future research should further explore these findings to develop more inclusive, effective, and sustainable volunteer programs that empower older adults to contribute meaningfully to their communities.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File, 61 KB - aging_v8i1e71810_app1.docx]

References

1. Population in brief. National Population and Talent Division. URL: https://www.population.gov.sg/media-centre/articles/population-in-brief-2023-key-trends/ [accessed 2024-11-22]
2. Mental health of older adults. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-12-02]
3. Institute of Mental Health. Well-being of the Singapore Elderly (WiSE) 2021 Study. URL: https://corp.nhg.com.sg/Media%20Releases/[illegible].pdf [accessed 2025-03-19]
4. Desai R, Tsipa A, Fearn C, et al. Suicide and dementia: a systematic review and meta-analysis of prevalence and risk factors. Ageing Res Rev 2024 Sep;100:102445. [doi: 10.1016/j.arr.2024.102445] [Medline: 39127443]
5. Low LL, Kwan YH, Ko MSM, et al. Epidemiologic characteristics of multimorbidity and sociodemographic factors associated with multimorbidity in a rapidly aging Asian country. JAMA Netw Open 2019 Nov 1;2(11):e1915245. [doi: 10.1001/jamanetworkopen.2019.15245] [Medline: 31722030]
6. Oh DJ, Han JW, Bae JB, et al. Chronic subsyndromal depression and risk of dementia in older adults. Aust N Z J Psychiatry 2021 Aug;55(8):809-816. [doi: 10.1177/0004867420972763] [Medline: 33198490]
7. Teo RH, Cheng WH, Cheng LJ, Lau Y, Lau ST. Global prevalence of social isolation among community-dwelling older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 Apr;107:104904. [doi: 10.1016/j.archger.2022.104904] [Medline: 36563614]
8. Lutz J, Van Orden KA. Sadness and worry in older adults: differentiating psychiatric illness from normative distress. Med Clin North Am 2020 Sep;104(5):843-854. [doi: 10.1016/j.mcna.2020.05.001] [Medline: 32773049]
9. National population health survey. Ministry of Health. URL: https://www.moh.gov.sg/others/resources-and-statistics/nphs-2023 [accessed 2024-12-02]
10. Volunteer with older people. Age UK. URL: https://www.ageuk.org.uk/get-involved/volunteer/ [accessed 2024-12-11]
11. Bringing out the best of America. AmeriCorps. URL: https://americorps.gov/about/our-impact [accessed 2024-12-11]
12. Kim ES, Whillans AV, Lee MT, Chen Y, VanderWeele TJ. Volunteering and subsequent health and well-being in older adults: an outcome-wide longitudinal approach. Am J Prev Med 2020 Aug;59(2):176-186. [doi: 10.1016/j.amepre.2020.03.004] [Medline: 32536452]
13. Volunteer opportunities. Lions Befrienders Service Association. URL: https://www.lionsbefrienders.org.sg/volunteer-opportunities/ [accessed 2024-12-11]
14. Morrow-Howell N, Hinterlong J, Rozario PA, Tang F. Effects of volunteering on the well-being of older adults. J Gerontol 2003 May;58(3):S137-S145. [doi: 10.1093/geronb/58.3.S137]
15. Singapore cares - building a caring and inclusive society. SG Cares. URL: https://www.sg/sgcares [accessed 2024-12-11]
16. Pinto da Costa M, Aviron P. Comparing the views of mental health professionals and volunteers on volunteering in mental health care in the UK. Discov Health Systems 2023 Dec 5;2(1):46. [doi: 10.1007/s44250-023-00056-4]

17. Pinto da Costa M, Oliveira J. Views on volunteering in mental health: a focus group study with mental health professionals and volunteers in Portugal. Discov Ment Health 2023 Jun 22;3(1):12. [doi: 10.1007/s44192-023-00038-1] [Medline: 37861821]
18. Brymer M, Jacobs A, Layne C, et al. Psychological First Aid: Field Operations Guide, 2nd edition: National Child Traumatic Stress Network and National Center for PTSD; 2006.
19. Vernberg EM, Steinberg AM, Jacobs AK, et al. Innovations in disaster mental health: psychological first aid. Prof Psychol: Res Pract 2008;39(4):381-388. [doi: 10.1037/a0012663]
20. Kelly EL, Fenwick KM, Brekke JS, Novaco RW. Sources of social support after patient assault as related to staff well-being. J Interpers Violence 2021 Jan;36(1-2):NP1003-NP1028. [doi: 10.1177/0886260517738779] [Medline: 29294965]
21. Pratt SD, Jachna BR. Care of the clinician after an adverse event. Int J Obstet Anesth 2015 Feb;24(1):54-63. [doi: 10.1016/j.ijoa.2014.10.001] [Medline: 25499810]
22. Wang L, Norman I, Xiao T, Li Y, Leamy M. Psychological first aid training: a scoping review of its application, outcomes and implementation. Int J Environ Res Public Health 2021 Apr 26;18(9):4594. [doi: 10.3390/ijerph18094594] [Medline: 33926108]
23. Bakes-Denman L, Mansfield Y, Meehan T. Supporting mental health staff following exposure to occupational violence - staff perceptions of “peer” support. Int J Ment Health Nurs 2021 Feb;30(1):158-166. [doi: 10.1111/inm.12767] [Medline: 33241642]
24. de Freitas Girardi J, Miconi D, Lyke C, Rousseau C. Creative expression workshops as Psychological First Aid (PFA) for asylum-seeking children: an exploratory study in temporary shelters in Montreal. Clin Child Psychol Psychiatry 2020 Apr;25(2):483-493. [doi: 10.1177/1359104519891760] [Medline: 31814432]
25. Schafer A, Snider L, Sammour R. A reflective learning report about the implementation and impacts of Psychological First Aid (PFA) in Gaza. Disaster Health 2016;3(1):1-10. [doi: 10.1080/21665044.2015.1110292] [Medline: 28229011]
26. Hartanto A, Sim L, Lee D, Majeed NM, Yong JC. Cultural contexts differentially shape parents' loneliness and wellbeing during the empty nest period. Commun Psychol 2024 Nov 3;2(1):105. [doi: 10.1038/s44271-024-00156-8] [Medline: 39489810]
27. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
28. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
29. Launch of Psychological First Aid (PFA) training for SGsecure responders. Ministry of Home Affairs. URL: https://www.mha.gov.sg/mediaroom/speeches/launch-of-psychological-first-aid-training-for-sgsecure-responders-closing-remarks-by-ms-sun-xueling-minister-of-state-ministry-of-home-affairs-ministry-of-social-and-family-development/ [accessed 2024-11-25]
30. Ng TKS, Feng L, Fam J, et al. Mindfulness Awareness Practice (MAP) to prevent dementia in older adults with mild cognitive impairment: protocol of a randomized controlled trial and implementation outcomes. Int J Environ Res Public Health 2021 Sep 28;18(19):10205. [doi: 10.3390/ijerph181910205] [Medline: 34639513]
31. Ruzek JI, Brymer MJ, Jacobs AK, Layne CM, Vernberg EM, Watson PJ. Psychological first aid. J Ment Health Couns 2007 Jan 1;29(1):17-49. [doi: 10.17744/mehc.29.1.5racqxjueafabgwp]
32. Cook A, Keyte R, Sprawson I, Matharu A, Mantzios M. Mental health first aid experiences: a qualitative investigation into the emotional impact of mental health first aid responsibilities and the significance of self-compassion. SN Soc Sci 2024 Nov 9;4(9):171. [doi: 10.1007/s43545-024-00962-1]
33. Horn R, O'May F, Esliker R, et al. The myth of the 1-day training: the effectiveness of psychosocial support capacity-building during the Ebola outbreak in West Africa. Glob Ment Health (Camb) 2019;6:e5. [doi: 10.1017/gmh.2019.2] [Medline: 31143466]
34. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Cypress BS. Rigor or reliability and validity in qualitative research: perspectives, strategies, reconceptualization, and recommendations. Dimens Crit Care Nurs 2017;36(4):253-263. [doi: 10.1097/DCC.0000000000000253] [Medline: 28570380]
37. Teh YY, Lek E. Culture and reflexivity: systemic journeys with a British Chinese family. J Fam Ther 2018 Nov;40(4):520-536. [doi: 10.1111/1467-6427.12205]
38. Buetow S. Apophenia, unconscious bias and reflexivity in nursing qualitative research. Int J Nurs Stud 2019 Jan;89:8-13. [doi: 10.1016/j.ijnurstu.2018.09.013] [Medline: 30316055]
39. Rocha ZL, Yeoh BSA. Measuring race, mixed race, and multiracialism in singapore. In: Rocha ZL, Aspinall PJ, editors. The Palgrave International Handbook of Mixed Racial and Ethnic Classification: Springer International Publishing; 2020:629-647.

40. Hotta K, Ishimaru M. The meaning in group volunteering and its influence on the well-being of community-dwelling older volunteers in Japan: a qualitative study. Arch Gerontol Geriatr Plus 2024 Dec;1(4):100086. [doi: 10.1016/j.aggp.2024.100086]
41. Same A, McBride H, Liddelow C, Mullan B, Harris C. Motivations for volunteering time with older adults: a qualitative study. PLOS ONE 2020;15(5):e0232718. [doi: 10.1371/journal.pone.0232718] [Medline: 32365081]
42. Withall J, Thompson JL, Fox KR, et al. Participant and public involvement in refining a peer-volunteering active aging intervention: project ACE (active, connected, engaged). Gerontologist 2018 Mar 19;58(2):362-375. [doi: 10.1093/geront/gnw148] [Medline: 27927733]
43. Barlow J, Hainsworth J. Volunteerism among older people with arthritis. Ageing Soc 2001 Mar;21(2):203-217. [doi: 10.1017/S0144686X01008145]
44. Okun MA, Michel J. Sense of community and being a volunteer among the young-old. J Appl Gerontol 2006 Apr;25(2):173-188. [doi: 10.1177/0733464806286710]
45. Welzel C. The Asian values thesis revisited: evidence from the world values surveys. Jpn J Political Sci 2011 Apr;12(1):1-31. [doi: 10.1017/S1468109910000277]
46. Hansen T, Slagsvold B. An “Army of Volunteers”? engagement, motivation, and barriers to volunteering among the baby boomers. J Gerontol Soc Work 2020;63(4):335-353. [doi: 10.1080/01634372.2020.1758269] [Medline: 32364438]
47. Serrat-Graboleda E, González-Carrasco M, Casas Aznar F, Malo Cerrato S, Cámara Liebana D, Roqueta-Vall-Llosera M. Factors favoring and hindering volunteering by older adults and their relationship with subjective well-being: a mixed-method approach. Int J Environ Res Public Health 2021 Jun 22;18(13):6704. [doi: 10.3390/ijerph18136704] [Medline: 34206359]
48. Best DL, Puzio AR. Gender and culture. In: Matsumoto D, Hwang HC, editors. The Handbook of Culture and Psychology: Oxford Academic; 2019:235-291. [doi: 10.1093/oso/9780190679743.003.0009]
49. Seow D. The role of gender, patriarchy and culture in the asian female travel experience. : Bournemouth University; 2020.
50. Hambrick EP, Rubens SL, Vernberg EM, Jacobs AK, Kanine RM. Towards successful dissemination of psychological first aid: a study of provider training preferences. J Behav Health Serv Res 2014 Apr;41(2):203-215. [doi: 10.1007/s11414-013-9362-y]
51. Ahmad NA, Abd Rauf MF, Mohd Zaid NN, Zainal A, Tengku Shahdan TS, Abdul Razak FH. Effectiveness of instructional strategies designed for older adults in learning digital technologies: a systematic literature review. SN Comput Sci 2022;3(2):130. [doi: 10.1007/s42979-022-01016-0] [Medline: 35039803]
52. Shorey S, Kua EH, Tam W, et al. “Where-There-Is-No-Psychiatrist Integrated Personal Therapy” among community-dwelling older adults: a randomized pilot study. Int J Environ Res Public Health 2021 Sep 9;18(18):9514. [doi: 10.3390/ijerph18189514] [Medline: 34574438]
53. Terry J. Experiences of instructors delivering the mental health first aid training programme: a descriptive qualitative study. J Psychiatr Ment Health Nurs 2010 Sep;17(7):594-602. [doi: 10.1111/j.1365-2850.2010.01557.x] [Medline: 20712682]
54. Schafer A, Snider L, van Ommeren M. Psychological first aid pilot: Haiti emergency response. Intervention (Amstelveen) 2010 Nov;8(3):245-254. [doi: 10.1097/WTF.0b013e32834134cb]
55. Semlitz L, Ogiwara K, Weissbecker I, et al. Psychological first aid training after Japan’s triple disaster: changes in perceived self competency. Int J Emerg Ment Health 2013;15(3):181-196. [Medline: 24558747]
56. Choi YJ. Psychological first-aid experiences of disaster health care workers: a qualitative analysis. Disaster Med Public Health Prep 2020 Aug;14(4):433-436. [doi: 10.1017/dmp.2019.87] [Medline: 31587684]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
PFA: Psychological First Aid

Edited by Y Jiang; submitted 27.01.25; peer-reviewed by K Machin, L Perrier; revised version received 02.04.25; accepted 07.04.25; published 13.06.25.

Please cite as:
Shorey S, Seow JSL, Mahendran R, Wee GH, Goh LG, Kua EH
Peer Volunteers’ Journeys Through Training and Engagement in Older Adult Communities: Descriptive Qualitative Study
JMIR Aging 2025;8:e71810
URL: https://aging.jmir.org/2025/1/e71810
doi:10.2196/71810

Frequency of Electronic Personal Health Record Use in US Older Adults: Cross-Sectional Study of a National Survey

Lavlin Agrawal[1*], PhD; Richelle Oakley DaSouza[1*], PhD; Pavankumar Mulgund[2*], PhD; Pankaj Chaudhary[1*], PhD

[1]Department of Business Information Systems and Analytics, Willie A. Deese College of Business and Economics, North Carolina Agricultural and Technical State University, 1601 E. Market Street, Greensboro, NC, United States

[2]Department of Management Information Systems, Fogelman College of Business & Economics, University of Memphis, Memphis, United States

[*]all authors contributed equally

Corresponding Author:
Lavlin Agrawal, PhD
Department of Business Information Systems and Analytics, Willie A. Deese College of Business and Economics, North Carolina Agricultural and Technical State University, 1601 E. Market Street, Greensboro, NC, United States

Abstract

Background: Electronic personal health records (ePHRs) hold significant potential to improve health management for older adults by enhancing access to medical information and facilitating communication with health care providers. However, usage remains low among individuals aged 65 and older. While existing research has identified barriers such as low self-efficacy, limited digital literacy, and usability challenges, the specific factors influencing the use of ePHRs among older adults are not yet fully understood.

Objective: This study integrates the Aging and Technology framework with the Patient Technology Acceptance Model to examine key predictors of ePHR use among older adults, including age, education, issue involvement, performance expectancy, effort expectancy, and self-efficacy, while controlling for demographic factors such as gender, race, and income.

Methods: This study utilizes data from the Health Information National Trends Survey (HINTS 5 cycle 3), which includes 532 respondents representing 13,136,180 US adults aged 65 years and older, after applying survey weights. Structural equation modeling was used to analyze the factors influencing the frequency of ePHR access over the past 12 months.

Results: The findings indicate that older adults with higher self-efficacy used ePHRs more frequently. Additionally, issue involvement, performance expectancy, and effort expectancy were positively associated with ePHR use. Notably, self-efficacy partially mediated the relationship between age and the frequency of ePHR use.

Conclusions: These findings suggest that enhancing self-efficacy, improving usability, and increasing the perceived benefits of ePHRs are critical for boosting usage among older adults. The study underscores the need for targeted interventions to support older users, simplify digital interfaces, and provide accessible educational resources, ultimately contributing to better health outcomes and an improved quality of life for older adults.

(JMIR Aging 2025;8:e71460) doi:10.2196/71460

KEYWORDS

electronic personal health record; aging population; self-efficacy; technology adoption; health IT; patient engagement

Introduction

Background

Health information technologies (HITs), such as electronic health records (EHRs) and electronic medical records (EMRs), are pivotal tools in modernizing health care by enhancing efficiency and improving the quality of care [1-3]. Beyond these general applications, HIT has the potential to address the unique medical needs of the older population. Extensive research demonstrates that HIT use can significantly improve health outcomes for older adults [4], particularly through tools such as EMRs and electronic personal health records (ePHRs). For example, EMRs in hospital settings help health care professionals identify at-risk patients, such as those prone to falls, by analyzing factors such as age and medication use [5]. Similarly, ePHRs empower older adults by providing easier access to health information, fostering improved communication with health care providers, and encouraging active engagement in personal health care management [6].

Luo et al [7] highlighted that, among older adults, the use of ePHRs was positively associated with the ease of understanding health information. ePHRs enable patients to track their health history to better understand the progress or deterioration of their health conditions. Patients with access to their health information experience increased medical data transparency, which reduces medical errors, improves trust in care providers, and enhances

patient satisfaction [8]. As such, in the United States, the adoption of ePHRs has steadily increased. This growth has been supported by policies such as the Health Information Technology for Economic and Clinical Health Act of 2009 [9], which incentivized health care providers to adopt EHR systems through financial rewards, indirectly creating an infrastructure supportive of ePHRs. The subsequent Meaningful Use Program set specific objectives for health care providers to demonstrate effective use of EHRs, including patient-access capabilities. Further support came from the 21st Century Cures Act of 2016 [10], which enhanced patient access to electronic health information and prohibited information blocking, thereby facilitating smoother integration and utilization of ePHRs.

While the benefits of HIT for older persons are well-documented, the adoption and sustained use of these technologies face notable barriers, such as lower levels of technology usage among older adults [11]. Recent data indicate that approximately 4 in 10 individuals aged 65 years and older utilize ePHRs, a rate significantly lower than that of younger populations [12]. Furthermore, a previous study reported that approximately 54% of US adults have been offered access to their online medical records, and among those offered, 57% accessed them at least once, translating to about 31% of the adult population accessing their online medical records [13]. However, specific statistics for individuals aged 65 years and older are not provided in this source. Given that older adults are generally less likely to use the internet, it is reasonable to infer that their rates of ePHR usage may be lower than those of younger populations [12].

Further, researchers have found that older age is negatively associated with HIT [14]. Reduced cognitive and motor abilities, common among older adults, often complicate interactions with modern digital applications [15,16]. van der Vaart et al [8] discussed how older age is associated with the nonuse of web-based portals. Czaja and Lee [17] argued that although older adults are willing to use technology, many report usability problems with existing systems, which may, in part, be due to the cognitive and perceptual demands placed on the user. Their findings show that self-efficacy is an important predictor of general use of technology and that people with lower self-efficacy are less likely to use technology. However, individual-level barriers such as confidence and usability are only part of the picture. A growing body of research suggests that technologies are often not designed with the specific needs of older adults in mind [18,19]. Studies have also shown that older patients may not be offered digital health tools for various reasons, including—but not limited to—age-related stereotypes held by health care professionals or concerns about overburdening older patients [20,21]. Together, these individual and environmental barriers create a complex landscape that can inhibit older adults' use of health technologies.

Despite advancements in technology, significant gaps remain in understanding the factors that influence the regular use of ePHRs among older adults. Existing research has focused on general barriers and facilitators, but limited knowledge exists about the interplay between these factors and how they shape the frequency of ePHR usage. This knowledge gap calls for a deeper examination of the factors affecting ePHR use in older populations. Additionally, limited electronic health literacy remains a critical challenge that hinders ePHR use in older populations [22]. However, the potential benefits cannot be overlooked. ePHRs offer older adults opportunities to coordinate their health care by sharing critical health information with providers and other stakeholders, an approach particularly crucial for managing chronic conditions that require multifaceted treatment strategies [7].

There are still unanswered questions about how older adults use ePHRs, particularly regarding the specific factors that facilitate or hinder ePHR use and the extent to which these factors influence it. To address this gap, this study explores the following research questions:

- What are the key factors influencing the use of ePHRs among older adults?
- How do age and education mediate the impact of self-efficacy on the frequency of ePHR use in older populations?
- What roles do facilitators such as issue involvement, performance expectancy, and effort expectancy of ePHRs play in encouraging regular use?

To answer these questions, we build on established theoretical frameworks such as the Patient Technology Acceptance Model (PTAM) [23], which provides insights into technology adoption and use processes. We extend the PTAM by incorporating factors specifically relevant to older persons, such as usability challenges, limited education, and issue involvement, to develop a comprehensive understanding of the dynamics at play. This study aims to contribute to the literature by offering empirical evidence and actionable insights into how ePHRs can be effectively designed and implemented to support the unique health care needs of aging populations.

Hypotheses and Proposed Model

Overview

This study aims to deepen the understanding of the frequency of PHR use among the older population by addressing critical individual factors that influence usage. A nuanced understanding of these factors is essential for improving ePHR-related outcomes, such as patient empowerment, health care accessibility, and overall health management. Specifically, this research investigates the determinants influencing the frequency of ePHR use among individuals aged 65 years and older.

Based on the existing literature, we contend that theoretical frameworks grounded in the Technology Acceptance Model (TAM) [24], including the PTAM [23], are particularly suitable for investigating frequency of use, aligning well with our study's specific outcomes. Previous research applying the TAM has effectively examined ongoing usage rather than merely initial adoption. For instance, McCloskey [25] utilized the TAM to evaluate sustained electronic commerce (e-commerce) behaviors, specifically the frequency of online purchasing, and found that perceived usefulness directly influenced the frequency of technology use. Similarly, Lederer et al [26] applied the TAM to explore frequent usage of the World Wide Web, explicitly addressing ongoing use rather than initial acceptance. Moreover, Martín-García et al [27] extended the TAM framework to

examine older adults' frequency of using technological devices, demonstrating the applicability of TAM constructs beyond initial intentions to sustained technological engagement.

Building upon this foundation, we integrate Czaja et al's [28] conceptual model of aging and technology with the PTAM [23]. Our selection of the PTAM specifically addresses our research context—older patients managing multiple chronic conditions at home using digital health tools. Unlike broader acceptance models such as the TAM [24], the Unified Theory of Acceptance and Use of Technology (UTAUT) [29], and the Health Belief Model [30], PTAM explicitly considers patient-centric variables critical for older populations managing complex health conditions.

For example, while the TAM broadly emphasizes general technological acceptance through perceived ease of use and perceived usefulness, it lacks explicit patient-oriented or health outcome considerations [31]. Similarly, the UTAUT primarily addresses technology acceptance within general workplace contexts rather than the nuanced health care needs of older patients with chronic conditions [32]. The Health Belief Model likewise falls short by inadequately addressing the interactive dynamics unique to older patient populations [33].

By contrast, PTAM's incorporation of patient-specific constructs—such as perceived health improvements, patient-provider relationships, and health care contextual factors—offers a more precise and comprehensive understanding of older patients' sustained engagement with digital health interventions. This tailored approach provides richer theoretical insights and clearer practical implications, significantly enhancing explanatory power and applicability compared with general acceptance models.

Age

The utilization of ePHRs and patient portals exhibits notable age-related disparities. Despite being frequent consumers of health care services, older adults are significantly less likely to use these digital tools compared with younger individuals [34]. This gap in use can be attributed to age-related barriers, such as limited digital literacy and difficulties in navigating technological interfaces, which become increasingly pronounced with age [34,35].

A systematic review of patient acceptance of HIT underscored this trend, finding a negative correlation between age and HIT acceptance [14]. Similarly, Heart and Kalderon [36] provided compelling evidence that older adults demonstrate lower use of health-related information and communication technologies. Additionally, van der Vaart et al [8] highlighted that older individuals are generally less likely to use patient web portals. Based on these findings, we propose the following hypothesis:

- *H1a: Age is negatively associated with the frequency of ePHR use.*

Research has consistently shown that as individuals age, their electronic self-efficacy—the belief in their ability to effectively use technology—tends to decline. This decrease in self-efficacy is often attributed to a combination of factors, including reduced exposure to new technologies, cognitive decline, and negative stereotypes associated with older adults and technology use. For example, studies suggest that older adults are more likely to perceive technology as complex and overwhelming, which can lower their confidence in using digital tools such as EMRs or patient portals [28]. With age, physiological capacity and the ability to respond to environmental stresses decline, further contributing to the reduction in electronic self-efficacy [37]. Hence, we hypothesize:

- *H1b: Age is negatively associated with self-efficacy.*

Education

Smith et al [38] argued that individuals with higher levels of education are more likely to register for and utilize patient portals. Empirical evidence supports this claim, indicating that higher educational attainment is positively correlated with increased patient portal usage [39]. Educated individuals are more likely to engage with digital platforms for managing their health information, as they typically possess greater health literacy and familiarity with technology. Furthermore, a recent literature review highlights that lower educational levels act as a barrier to ePHR use [40]. Similarly, individuals with higher education are more likely to perceive the benefits of ePHRs, leading to higher usage rates [41]. Based on these findings, we propose the following hypothesis:

- *H2a: Education is positively associated with the frequency of ePHR use*.

Research consistently demonstrates that education and training play a pivotal role in enhancing individuals' electronic self-efficacy, particularly in the context of using digital tools such as patient portals and EMRs. Czaja et al [28] highlighted a positive association between education and self-efficacy, emphasizing the importance of foundational skills acquired through educational experiences. These skills, including problem-solving, critical thinking, and basic computer literacy, are critical for successfully navigating digital platforms [42]. Furthermore, education has been shown to empower individuals across all age groups to overcome barriers to technology use. Based on this, we propose the following hypothesis:

- *H2b: Education is positively associated with self-efficacy.*

Issue Involvement

Issue involvement in the health care domain refers to "how relevant a specific health issue is to a patient" [43,44]. A more involved patient typically has a severe health condition and frequently visits health care providers [45]. Ross et al [46] argued that issue involvement has a significant positive impact on the use of patient-accessible medical records. This finding is further confirmed by Angst and Agarwal [45] and Abdelhamid et al [43]. A more involved patient is more likely to use ePHRs, as they help patients better prepare for upcoming visits with physicians by enhancing their knowledge of their medical condition, increasing their sense of control, and allowing them to seek clarification about treatment. Hence, we hypothesize:

- *H3: Issue involvement is positively associated with the frequency of ePHR use.*

Performance Expectancy

Performance expectancy is defined as the degree to which a person feels that using a system will help them perform a job more efficiently [29]. In keeping with this understanding, we refer to performance expectancy as the degree to which the patient believes that using ePHRs helps them monitor their health. Venkatesh et al [29] theorized that performance expectancy drives the intention to use information systems. Several researchers have also identified performance expectancy as one of the critical predictors of eHealth acceptance and use [40,47-49]. Patient portals help improve patient engagement and empower individuals to access their health information anytime and anywhere [50]. Thus, we hypothesize:

- *H4: Performance expectancy is positively associated with the frequency of ePHR use.*

Effort Expectancy

Venkatesh et al [29] defined effort expectancy as the degree of comfort associated with system use. Consistent with Venkatesh et al [29], we define effort expectancy as the degree of ease associated with understanding the health information in the online medical record. Effort expectancy is another key variable that drives use intentions [46]. Many researchers have also identified effort expectancy as one of the critical predictors of health technology use [2,23,37,48]. Therefore, we also propose the following hypothesis:

- *H5: Effort expectancy is positively associated with the frequency of ePHR use.*

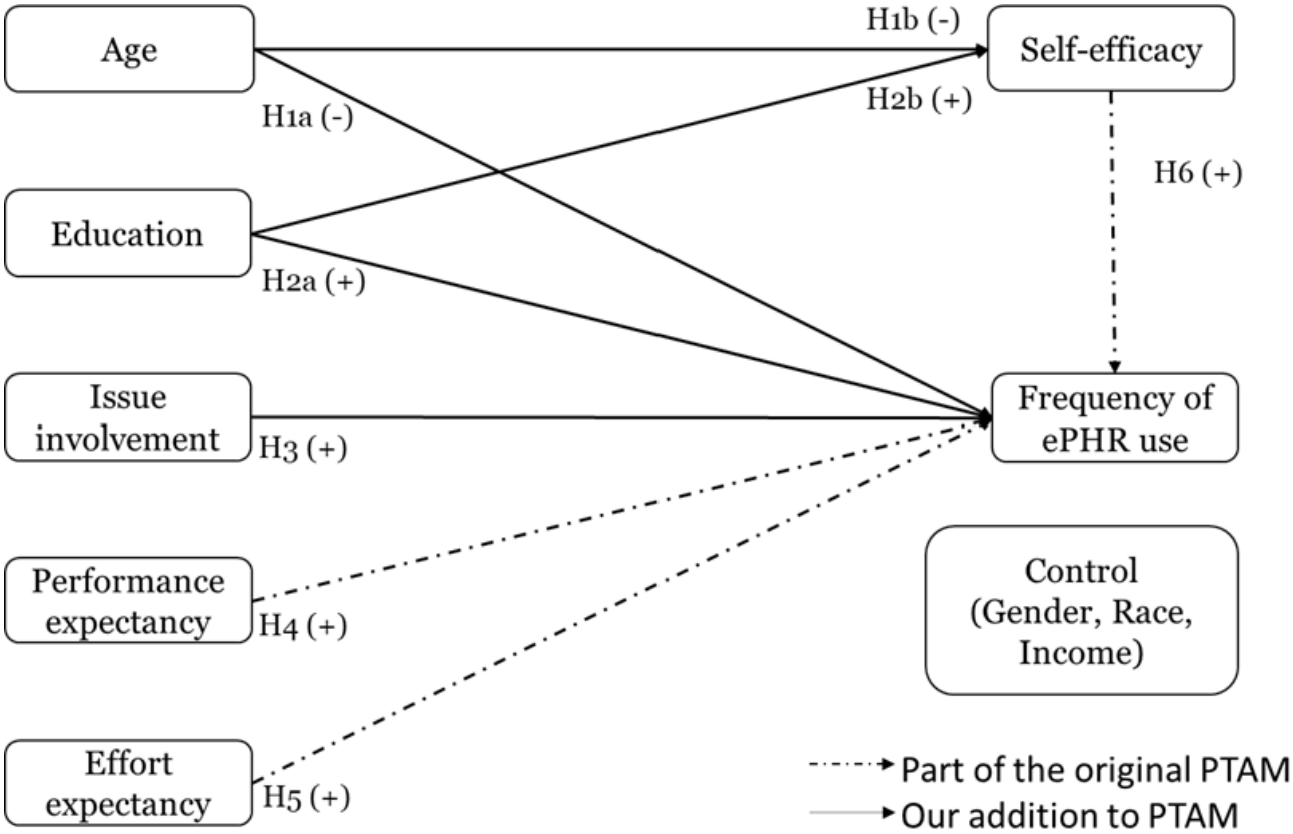
Self-Efficacy

Research has consistently demonstrated that electronic self-efficacy—the belief in one's capability to effectively use technology—declines with age due to factors such as reduced exposure to new technologies, cognitive decline, diminished physiological capacity, and negative stereotypes about older adults' technology use [28,37]. Consequently, older adults often perceive technology, including ePHRs, as complex and overwhelming, undermining their confidence in engaging with such digital tools. Therefore, we selected self-efficacy as a mediator to explicitly capture the psychological mechanism linking age with ePHR use. Rather than age directly limiting ePHR engagement, we posit that reduced electronic self-efficacy partially explains why older adults use ePHRs less frequently. This approach emphasizes the central role of self-efficacy in technology use among aging populations [36,51]. Hence, we hypothesize:

- *H6: Self-efficacy is positively associated with the frequency of ePHR use.*

Our proposed model is depicted in Figure 1. Additionally, we control for demographic factors such as gender, race, and income to account for their potential influence on the frequency of ePHR use.

Figure 1. The research model. ePHR: electronic personal health record; PTAM: Patient Technology Acceptance Model.



Methods

Data Source

This study used publicly available cross-sectional data collected in 2019 by the National Cancer Institute (Health Information National Trends Survey [HINTS] 5 cycle 3). The cross-sectional analysis for this study was limited to a subset of the original dataset and included participants aged 65 years or older. The total number of individuals aged 65 years or older was 1961. Of these, 1234 had not used an ePHR at least once in the past 12 months and were therefore excluded from the analysis. Further, 195 responses were removed due to missing values for key variables. The final sample consisted of 532 respondents, representing 13,136,180 US adults aged 65 and older when survey weights were applied.

Ethical Considerations

The HINTS 5 survey, conducted with the general population, underwent expedited review and received approval from the Westat Institutional Review Board on March 28, 2016 (project number 6048.14). This analysis used deidentified, publicly available data from HINTS, which did not constitute human research as defined by 45 CFR 46.102 and, therefore, did not require Institutional Review Board review.

Measurements

In this study, the primary dependent variable, “frequency of ePHR use,” is operationalized using a single-item measure assessing how frequently patients accessed their online medical records within the past 12 months. The response options are categorized as follows: 1-2 times, 3-5 times, 6-9 times, and 10 or more times. Individuals who reported not accessing their online medical records during the past year were excluded from the analysis, as critical variables pertinent to these respondents were not captured.

Performance expectancy, representing the perceived usefulness of online medical records for health monitoring, is measured using a 5-point Likert scale, ranging from “5=very useful” to “1=don’t use.” Effort expectancy, reflecting the perceived effort necessary to understand health information in online medical records, is assessed using a 4-point Likert scale ranging from “4=very easy” to “1=very difficult.” Issue involvement captures an individual’s engagement with personal health management and is operationalized as the number of interactions a respondent has had with health care providers over the past year. Self-efficacy in accessing electronic health records is assessed through a binary-choice question indicating whether respondents have utilized any electronic method to access their medical records during the past 12 months.

Single-item measures are used for several constructs in this study. Such measures are considered acceptable when questions are clear, unambiguous, and not prone to multiple interpretations [52]. Additionally, single-item measures are prevalent and widely validated within information systems research, particularly in studies applying structural equation modeling (SEM) in health care contexts [43,44].

Gender, race, and income are incorporated as control variables, consistent with prior studies [43,44]. For comprehensive details on questionnaire items, scales, and specific variable operationalizations, please refer to Multimedia Appendix 1.

Statistical Analysis

This study uses SEM to conduct a path analysis. Although SEM is predominantly used to model latent variables, it is also commonly applied to conduct a path analysis in mediation models. We examine 2 mediating relationships: first, self-efficacy mediates the relationship between age and the frequency of ePHR use; second, self-efficacy also mediates the relationship between education and the frequency of ePHR use. Accordingly, we use SEM to test the model, consistent with prior research [53,54]. We used SEM with weighted least squares mean and variance adjusted estimation to test the hypotheses. Weighted least squares mean and variance adjusted is well-suited for models with ordinal outcome variables [55,56]. The analysis was conducted in R (R Foundation) using the “lavaan.survey” package. We incorporated HINTS-supplied survey weights and used jackknife variance estimation techniques to account for the complex HINTS sampling design and to calculate nationally representative estimates [57].

Results

Descriptive Statistics

Table 1 shows the descriptive statistics of the survey respondents. The survey included questions about the frequency of participants’ ePHR use, as well as questions related to the model variables, including performance expectancy, effort expectancy, issue involvement, self-efficacy, age, education, gender, race, and income.

Table . Descriptive statistics.

Demographic characteristics	Sample size (N=532)[a]	With survey weights (N=13,136,180)[b]
Gender		
Male	249 (46.8)	5,729,371
Female	283 (53.2)	7,406,809
Race		
White	458 (86.1)	11,961,412
Non-White	74 (13.9)	1,174,768
Education		
Less than high school	7 (1.3)	278,445
High school or higher	525 (98.7)	12,857,735
Income (US $)		
Less than 20,000	57 (10.7)	1,191,896
20,000 to <35,000	68 (12.8)	2,009,787
35,000 to <50,000	86 (16.2)	2,100,148
50,000 to <75,000	122 (22.9)	3,110,545
75,000 or higher	199 (37.4)	4,723,804
Age (years)		
Range	65-97	65-97
Mean (SD)	71.68 (5.66)	71.90 (6.09)

[a]The sample size for gender, race, education, and income is presented as n (%).

[b]Data are presented as n for gender, race, education, and income.

Reliability and Validity

Table 2 presents the correlations between all variables of interest. Correlation coefficients are important because high correlations among independent variables may indicate multicollinearity, which can introduce bias into the model results. Multicollinearity is not a concern in this analysis, as all correlations fall within the acceptable threshold of 0.6 [58], with the highest correlation being 0.37 between performance expectancy and effort expectancy. Table 2 also provides the means and SDs for the principal variables.

Table . Correlation matrix.

Correlation	Mean (SD)	Performance expectancy	Effort expectancy	Issue involvement	Self-efficacy
Performance expectancy	4.15 (1.03867)	1.00	—[a]	—	—
Effort expectancy	3.30 (0.67763)	0.37	1.00	—	—
Issue involvement	3.80 (1.49733)	−0.01	−0.14	1.00	—
Self-efficacy	0.82 (0.38126)	0.17	0.06	0.09	1.00

[a]Not applicable.

Common Method Variance

As data are self-reported and collected through a single survey, they may suffer from common method variance (CMV), which hampers the relationship between the variables [59]. We used the marker variable technique [60] to check if the data are suffering from CMV. Marker variable is a variable that is theoretically unrelated to 1 or more of the principal variables measured in the study and typically has a low correlation with the central variables.

The theoretically unrelated construct “morning person or night person” was used as a marker variable. The correlations between the marker variable morning person or night person and other principal variables were below the threshold of 0.1 (performance expectancy: −0.08; effort expectancy: −0.08; self-efficacy: 0.03) [60], except for issue involvement (0.17). The low correlation of the marker variable with the principal variables in the model indicates that CMV is not a problem.

Data Analysis

As the National Cancer Institute administered both a paper-based questionnaire and an online questionnaire to survey participants, we first assessed whether the mode of survey introduced any biases. We regressed the dependent variable, "frequency of ePHR use," on the mode of survey and found that the relationship between the 2 was nonsignificant (*P*=.48). Thus, the mode of survey administration did not introduce any bias in the main outcome variable, that is, "frequency of ePHR use."

The SEM results are presented in Table 3. The overall fit statistics (χ^2_{17}=13.230, *P*=.004; Comparative Fit Index=0.950, Tucker-Lewis Index=0.715, root-mean-square error of approximation=0.080, standardized root-mean-square residual=0.022) indicate a good model fit [61]. Table 4 presents the results of the mediation analysis for self-efficacy with age and education.

The results show that age is negatively associated with self-efficacy (β=−.9890, *P*=.03), suggesting that as age increases, self-efficacy decreases, supporting H1b. However, no significant (*P*=.85) relationship was found between self-efficacy and education (H2b) in our research context.

Further, the results show that age is positively associated with the frequency of ePHR use (β=1.4950, *P*=.03), suggesting that as age increases, ePHR use also increases. This is the opposite of what we hypothesized for H1a. This counterintuitive trend may reflect increased health care needs among older adults, leading to greater engagement with ePHRs for managing chronic conditions, medication schedules, and communication with physicians. No significant (*P*=.88) relationship was found between education and the frequency of ePHR use (H2a) in our research context. H3 proposed a positive relationship between issue involvement and the frequency of ePHR use. The path coefficient was positive and statistically significant (β=.2560, *P*<.001), suggesting that higher issue involvement leads to higher ePHR use.

Our analysis also revealed a significant positive relationship between performance expectancy and the frequency of ePHR use (β=.2470, *P*<.001), as well as between effort expectancy and the frequency of ePHR use (β=.1850, *P*=.04). These findings suggest that higher performance expectancy and higher effort expectancy both lead to increased ePHR use, supporting H4 and H5. We also found that self-efficacy is positively associated with the frequency of ePHR use (β=.4990, *P*<.001), supporting H6.

Further analysis confirms that self-efficacy partially mediates the relationship between age and the frequency of ePHR use among older adults, as evidenced by the statistical significance of both indirect (*P*=.03) and direct effects (*P*=.03). Specifically, self-efficacy accounts for 49.2% (−0.493/1.002) of the total effect, underscoring its substantial role as a mediator in this relationship. However, we did not find any statistically significant mediating effect of self-efficacy on the relationship between education and the frequency of ePHR use.

Table . Results of structural equation modeling.

Dependent variable and hypothesis	Variables	Estimate	95% CI	*P* value	Significant
Self-efficacy					
H1b	Log (age)	−0.9890	−1.8870 to −0.0900	.03	Yes
H2b	High school or more	−0.0360	−0.4070 to 0.3350	.85	No
Control	Female	−0.0190	−0.1300 to 0.0910	.73	No
Control	White	−0.0040	−0.1460 to 0.1370	.95	No
Control	Income	0.0120	−0.0200 to 0.0430	.47	No
Dependent variable: frequency of electronic patient health record use					
H1a	Log (age)	1.4950	0.1460 to 2.8440	.03	Yes
H2a	High school or more	0.0320	−0.3770 to 0.4410	.88	No
H3	Issue involvement	0.2560	0.1810 to 0.3300	<.001	Yes
H4	Performance expectancy	0.2470	0.1340 to 0.3600	<.001	Yes
H5	Effort expectancy	0.1850	0.0070 to 0.3640	.04	Yes
H6	Self-efficacy	0.4990	0.3260 to 0.6730	<.001	Yes
Control	Female	0.1970	−0.0410 to 0.4340	.10	No
Control	White	0.0590	−0.3050 to 0.4220	.75	No
Control	Income	−0.0300	−0.1290 to 0.0680	.54	No

Table . Mediation results of structural equation modeling.

Variables	Estimate	95% CI	*P* value	Significant
Log (age)				
Indirect through self-efficacy	−0.4930	−0.9410 to −0.0460	.03	Yes
Direct	1.4950	0.1460 to 2.8440	.03	Yes
Total	1.0020	−0.3170 to 2.3200	.14	No
High school or more				
Indirect through self-efficacy	−0.0180	−0.2040 to 0.1680	.85	No
Direct	0.0320	−0.3770 to 0.4410	.88	No
Total	0.0140	−0.4860 to 0.5140	.96	No

Discussion

Determinants of ePHR Use

Our study aimed to update the PTAM by examining the impact of performance expectancy, effort expectancy, and self-efficacy on the frequency of ePHR Use. By using an integrated framework, this research provides actionable insights into the factors driving ePHR usage among older adults, incorporating variables that have not been widely explored in prior studies. This study makes both a significant theoretical and practical contribution.

Theoretical Implications

This study contributes to the literature on ePHR use by building on and extending existing theoretical frameworks, including the Aging and Technology framework [28] and the PTAM [23]. Prior research has predominantly highlighted barriers to ePHR use among older adults, such as reduced self-efficacy, usability challenges, and low digital literacy [15,17]. While these studies provide an important foundation, our findings challenge and refine these perspectives by demonstrating a positive association between age and ePHR use, with self-efficacy serving as a partial mediator in this relationship.

Contrary to earlier work suggesting older age is a barrier to HIT use [8,14], this study reveals that age positively influences ePHR use when health needs increase, as chronic illnesses and frequent health care interactions necessitate greater reliance on digital tools. This finding aligns with studies emphasizing the contextual nature of technology use, where health-related motivations can offset age-related challenges [7]. By demonstrating that older adults use ePHRs more frequently despite lower self-efficacy, this study extends theoretical models by incorporating health-related drivers, such as issue involvement, into the broader narrative of technology adoption.

Self-efficacy, identified as a strong predictor and mediator, validates and expands prior work in the field. Studies have consistently highlighted self-efficacy as a key determinant of technology use [62]. Our findings reinforce this and reveal that self-efficacy mediates the impact of both age and education on ePHR use. This underscores the critical role of psychological factors, particularly in populations facing cognitive and physical challenges, and aligns with theories from the PTAM [23] and the UTAUT [29].

The study also affirms the relevance of performance expectancy and effort expectancy as critical drivers of ePHR usage, consistent with the PTAM and UTAUT frameworks [47,48]. However, by identifying issue involvement as a significant predictor, this study contributes a novel dimension to the literature. This finding supports prior arguments that perceived relevance and personal health involvement significantly enhance engagement with health technologies [43,45].

Overall, this research advances theoretical understanding by integrating psychological, motivational, and contextual factors into established frameworks for HIT use. It bridges gaps in the existing literature by demonstrating that age, when considered alongside mediating factors such as self-efficacy and health involvement, can positively influence ePHR use. These insights offer a nuanced understanding of older adults' interactions with health technologies, offering a robust foundation for future research and practical interventions.

Practical Implications

Overview

The findings of this study underscore several actionable strategies for health care providers, policy makers, and technology developers to enhance ePHR usage among older adults. By addressing barriers and leveraging the facilitators identified, targeted interventions can empower older adults to better manage their health through ePHRs, contributing to improved health outcomes and broader system-wide benefits, as outlined below.

Building Self-Efficacy Through Tailored Training Programs

To enhance older adults' self-efficacy in using ePHRs, health care organizations can establish targeted training and education programs aligned with the recommendations of the Office of the National Coordinator for Health Information Technology [63]. These programs could include workshops in collaboration with senior community centers and local health care providers, focusing on essential tasks such as medication tracking, viewing test results, and scheduling appointments [64]. Additionally, peer-led mentoring initiatives, consistent with community-based

health literacy efforts promoted by Healthy People 2030, can encourage skill sharing among older adults [65].

Supporting Continuous Improvement and User-Centered Innovation

Continuous evaluation through user feedback mechanisms should be institutionalized to guide the ongoing improvement of ePHR technologies [66], particularly for older adults. Actively encouraging the participation of older adults in innovation processes can help health care organizations advance the user-centered innovation strategies promoted by federal health technology initiatives [63]. Collaboration among academia, health care providers, and technology developers can further drive innovation tailored to the evolving needs of older adults.

Promoting Health Involvement Through Guided Support

Health care providers can encourage older adults' active participation in managing their health through targeted education campaigns that highlight the practical benefits of ePHRs. Public health strategies that emphasize patient-provider communication, administrative ease, and improved health care outcomes can be strengthened by incorporating user testimonials and success stories [67]. Additionally, guided tutorials embedded within ePHRs can further support patient engagement, aligning with strategies recommended by the Agency for Healthcare Research and Quality for patient empowerment [68].

Implementing these recommendations, grounded in existing US health care policies and strategies, can significantly enhance ePHR adoption among older adults, help reduce health disparities, and promote a more efficient and patient-centered health care system.

Limitations

This study has several limitations that should be considered. First, the use of secondary data from the HINTS limited the analysis to variables available in the dataset. Important factors, such as detailed measures of digital literacy or prior technology experience, could not be examined. Further, this study included only individuals who reported using ePHRs at least once in the past 12 months, as key variables of interest were not captured for those who had not used ePHRs. Additionally, some variables in this study were measured using single-item scales, which may lack the robustness of multi-item measures. However, single-item measures are considered acceptable when the constructs are straightforward and unambiguous [52]. We acknowledge that single-item measures inherently do not allow for traditional assessments of internal consistency reliability, which presents a limitation in this study. This limitation necessitates cautious interpretation of the findings. Future research should aim to incorporate multi-item scales where feasible and conduct reliability analyses to ensure the stability and consistency of these measures.

Second, the reliance on self-reported data introduces the possibility of CMV, which could inflate relationships between variables [59]. To address this concern, the study applied the marker variable technique [53], which confirmed that CMV was not a significant issue. Third, the cross-sectional design of the study limits the ability to establish causal relationships among the variables. Longitudinal research is needed to explore how factors such as self-efficacy and performance expectancy influence the use of ePHRs over time. Fourth, although the use of survey weights enhances the representativeness of the sample for the United States, the relatively small sample size remains a limitation that future research with a more targeted approach could address. Additionally, we did not control for factors such as income, digital literacy, and prior technology experience, as these variables were not accessible in the dataset. The absence of these controls may limit the completeness of the study. Future research should incorporate these variables to provide a more comprehensive understanding of the frequency of HIT use among older adults. Finally, the findings are specific to the US context and may not fully generalize to other populations or health care systems. Despite these limitations, the study offers valuable insights and lays important groundwork for future research.

Conclusions

This study emphasizes the important role of ePHRs in empowering older adults to manage their health and maintain independence. By utilizing survey weights, the findings can be generalized to the broader US population, making them particularly relevant as the aging demographic continues to grow. Unlike previous research, this study reveals a positive relationship between age and the level of ePHR use, challenging the common assumption that older adults are reluctant to adopt health technologies.

Key factors driving ePHR usage are performance expectancy, effort expectancy, self-efficacy, and issue involvement, all of which offer actionable pathways for increasing the frequency of use among older adults. Raising awareness of the practical benefits and ease of use of ePHRs can encourage more frequent engagement. Additionally, addressing usability concerns and emphasizing the relevance of ePHRs to individual health needs can further promote their adoption. Frequent use of ePHRs empowers older adults to manage chronic conditions, access vital health information, and make informed decisions, ultimately enhancing their quality of life.

As health care systems strive to meet the challenges posed by an aging population, integrating ePHRs into routine care and ensuring equitable access can provide both social and economic benefits. This study lays the groundwork for targeted interventions aimed at bridging the digital divide and fostering a more inclusive, health-empowered society for older adults.

Acknowledgments

The authors thank the reviewers and editor for constructive feedback.

Disclaimer

Generative artificial intelligence tools such as Grammarly and ChatGPT were used to assist with language editing and clarity in certain sections of the manuscript. The article is an original work by the authors, who take full responsibility for its content.

Data Availability

The dataset used in this study is publicly available from the National Cancer Institute's Health Information National Trends Survey (HINTS) portal [69].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Operationalization of constructs.
[DOCX File, 22 KB - aging_v8i1e71460_app1.docx]

References

1. What are the advantages of electronic health records? Office of the National Coordinator for Health Information Technology. 2024. URL: https://www.healthit.gov/faq/what-are-advantages-electronic-health-records [accessed 2025-07-13]
2. Dutta B. What motivates people to adopt personal health record? Understanding the effect of innovation and context-specific threat on personal health record adoption intention. International Journal of Human–Computer Interaction 2024 Sep;40(17):4831-4844. [doi: 10.1080/10447318.2023.2223821]
3. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [doi: 10.7326/0003-4819-144-10-200605160-00125] [Medline: 16702590]
4. Davaridolatabadi N, Sadoughi F, Ahmadi M. Health information system to improve elderly health: study of health information system in us to setup standard for developing countries. Medical Gazette/Medicinski Glasnik 2013;18:50. [doi: 10.5937/medgla1350018D]
5. Weber V, White A, McIlvried R. An electronic medical record (EMR)-based intervention to reduce polypharmacy and falls in an ambulatory rural elderly population. J Gen Intern Med 2008 Apr;23(4):399-404. [doi: 10.1007/s11606-007-0482-z] [Medline: 18373136]
6. Perotti L, Stamm O, Dietrich M, Buchem I, Müller-Werdan U. The usability and user experience of an interactive e-learning platform to empower older adults when using electronic personal health records: an online intervention study. Universal Access in the Information Society 2025;24:1193-1208.
7. Luo RZ, Lu Q, Sun Y, Sun H, Liu YH, Zhao Y. Investigation on risk factors of chronic diseases among community residents: a study based on health management systems supported by mobile phones. Nurs Open 2023 Aug;10(8):5117-5128. [doi: 10.1002/nop2.1747] [Medline: 37071702]
8. van der Vaart R, Drossaert CHC, Taal E, Drossaers-Bakker KW, Vonkeman HE, van de Laar M. Impact of patient-accessible electronic medical records in rheumatology: use, satisfaction and effects on empowerment among patients. BMC Musculoskelet Disord 2014 Mar 26;15(1):102. [doi: 10.1186/1471-2474-15-102] [Medline: 24673997]
9. HITECH Act enforcement interim final rule. US Department of Health and Human Services. 2009. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/hitech-act-enforcement-interim-final-rule/index.html [accessed 2025-07-13]
10. 21st Century Cures Act. US Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act [accessed 2025-07-13]
11. An S, Cheung CF, Lo YT. Improving older adults' technology adoption on mobile map: a gamification approach. Int J Hum Comput Interact 2024:1-19. [doi: 10.1080/10447318.2024.2426914]
12. Share of those 65 and older who are tech users has grown in the past decade. Pew Research Center. 2022 Jan 13. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2025-07-13]
13. Trivedi N, Patel V, Johnson C, Chou WYS. Barriers to accessing online medical records in the United States. Am J Manag Care 2021 Jan;27(1):33-40. [doi: 10.37765/ajmc.2021.88575] [Medline: 33471460]
14. Or CKL, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560. [doi: 10.1197/jamia.M2888] [Medline: 19390112]
15. Taha J, Czaja SJ, Sharit J, Morrow DG. Factors affecting usage of a personal health record (PHR) to manage health. Psychol Aging 2013 Dec;28(4):1124-1139. [doi: 10.1037/a0033911] [Medline: 24364414]
16. Hourcade JP, Chrischilles EA, Gryzlak BM, et al. Design lessons for older adult personal health records software from older adults. Presented at: Universal Access in Human-Computer Interaction Users Diversity: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011; Jul 9-14, 2011; Orlando, FL p. 176-185. [doi: 10.1007/978-3-642-21663-3_19]

17. Czaja SJ, Lee CC. The impact of aging on access to technology. Univ Access Inf Soc 2007 Mar 19;5(4):341-349. [doi: 10.1007/s10209-006-0060-x]
18. Peine A, Neven L. From intervention to co-constitution: new directions in theorizing about aging and technology. Gerontologist 2019 Jan 9;59(1):15-21. [doi: 10.1093/geront/gny050] [Medline: 29850812]
19. Gallistl V, Rohner R, Seifert A, Wanka A. Configuring the older non-user: between research, policy and practice of digital exclusion. SI 2020;8(2):233-243. [doi: 10.17645/si.v8i2.2607]
20. Mannheim I, Wouters EJM, van Boekel LC, van Zaalen Y. Attitudes of health care professionals toward older adults' abilities to use digital technology: questionnaire study. J Med Internet Res 2021 Apr 21;23(4):e26232. [doi: 10.2196/26232] [Medline: 33881408]
21. Spann A, Stewart E. Barriers and facilitators of older people's mHealth usage: a qualitative review of older people's views. Human Technology 2018;14(3):264-296. [doi: 10.17011/ht/urn.201811224834]
22. Noblin AM, Wan TTH, Fottler M. The impact of health literacy on a patient's decision to adopt a personal health record. Perspect Health Inf Manag 2012;9(Fall):1-13. [Medline: 23209454]
23. Or CKL, Karsh BT, Severtson DJ, Burke LJ, Brown RL, Brennan PF. Factors affecting home care patients' acceptance of a web-based interactive self-management technology. J Am Med Inform Assoc 2011;18(1):51-59. [doi: 10.1136/jamia.2010.007336] [Medline: 21131605]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
25. McCloskey D. Evaluating electronic commerce acceptance with the Technology Acceptance Model. JCIS 2004;44(2):49-57. [doi: 10.1080/08874417.2004.11647566]
26. Lederer AL, Maupin DJ, Sena MP, Zhuang Y. The technology acceptance model and the World Wide Web. Decis Support Syst 2000 Oct;29(3):269-282. [doi: 10.1016/S0167-9236(00)00076-2]
27. Martín-García AV, Redolat R, Pinazo-Hernandis S. Factors influencing intention to technological use in older adults. The TAM model aplication. Res Aging 2022;44(7-8):573-588. [doi: 10.1177/01640275211063797] [Medline: 34962846]
28. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
29. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
30. Rosenstock IM. The Health Belief Model and preventive health behavior. Health Educ Monogr 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
31. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
32. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the Unified Theory of Acceptance and Use of Technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
33. Champion VL, Skinner CS. The health belief model. In: Health Behavior and Health Education: Theory, Research, and Practice: Jossey-Bass; 2008, Vol. 4:45-65.
34. Khadjesari Z, Houghton J, Brown TJ, Jopling H, Stevenson F, Lynch J. Contextual factors that impact the implementation of patient portals with a focus on older people in acute care hospitals: scoping review. JMIR Aging 2023;6:e31812. [doi: 10.2196/31812] [Medline: 36735321]
35. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024;24. [doi: 10.1186/s12889-024-18036-5]
36. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
37. Troen BR. The biology of aging. Mt Sinai J Med 2003 Jan;70(1):3-22. [Medline: 12516005]
38. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Jul;22(4):888-895. [doi: 10.1093/jamia/ocv025] [Medline: 25914099]
39. Dendere R, Slade C, Burton-Jones A, Sullivan C, Staib A, Janda M. Patient portals facilitating engagement with inpatient electronic medical records: a systematic review. J Med Internet Res 2019 Apr 11;21(4):e12779. [doi: 10.2196/12779] [Medline: 30973347]
40. Niazkhani Z, Toni E, Cheshmekaboodi M, Georgiou A, Pirnejad H. Barriers to patient, provider, and caregiver adoption and use of electronic personal health records in chronic care: a systematic review. BMC Med Inform Decis Mak 2020 Jul 8;20(1):153. [doi: 10.1186/s12911-020-01159-1] [Medline: 32641128]
41. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12. [doi: 10.1177/1357633X16674087] [Medline: 29320966]
42. van Deursen AJ, van Dijk JA. The digital divide shifts to differences in usage. New Media & Society 2014 May;16(3):507-526. [doi: 10.1177/1461444813487959]

43. Abdelhamid M, Gaia J, Sanders GL. Putting the focus back on the patient: how privacy concerns affect personal health information sharing intentions. J Med Internet Res 2017 Sep 13;19(9):e169. [doi: 10.2196/jmir.6877] [Medline: 28903895]
44. Agrawal L, Ndabu T, Mulgund P, Sharman R. Factors affecting the extent of patients' electronic medical record use: an empirical study focusing on system and patient characteristics. J Med Internet Res 2021 Oct 28;23(10):e30637. [doi: 10.2196/30637] [Medline: 34709181]
45. Angst C, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Q 2009;33(2):339. [doi: 10.2307/20650295]
46. Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin CT. Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 2005 May 24;7(2):e13. [doi: 10.2196/jmir.7.2.e13] [Medline: 15914460]
47. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers J, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103. [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
48. Tavares J, Oliveira T. Electronic health record portal adoption: a cross country analysis. BMC Med Inform Decis Mak 2017 Jul 5;17(1):97. [doi: 10.1186/s12911-017-0482-9] [Medline: 28679423]
49. Mukhopadhyay S, Basak R, Carpenter D, Reithel B. Patient use of online medical records: an application of technology acceptance framework. Information & Computer Security 2019;28(1):97-115 [FREE Full text] [doi: 10.1108/ICS-07-2019-0076]
50. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. J Am Med Inform Assoc 2011 Dec;18 Suppl 1(Suppl 1):i8-12. [doi: 10.1136/amiajnl-2011-000473] [Medline: 22071530]
51. Chen K, Chan AHS. A review of technology acceptance by older adults. Gerontechnology 2011;10(1). [doi: 10.4017/gt.2011.10.01.006.00]
52. Wanous JP, Reichers AE, Hudy MJ. Overall job satisfaction: how good are single-item measures? J Appl Psychol 1997 Apr;82(2):247-252. [doi: 10.1037/0021-9010.82.2.247] [Medline: 9109282]
53. Santibáñez-Andrade G, Castillo-Argüero S, Vega-Peña EV, Lindig-Cisneros R, Zavala-Hurtado JA. Structural equation modeling as a tool to develop conservation strategies using environmental indicators: the case of the forests of the Magdalena river basin in Mexico City. Ecol Indic 2015 Jul;54:124-136. [doi: 10.1016/j.ecolind.2015.02.022]
54. MacKinnon D, Cheong J, Pirlott A. Statistical Mediation Analysis: American Psychological Association; 2012.
55. DiStefano C, Morgan GB. A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. Struct Equ Modeling 2014 Jul 3;21(3):425-438. [doi: 10.1080/10705511.2014.915373]
56. Nguyen TQ, Webb-Vargas Y, Koning IM, Stuart EA. Causal mediation analysis with a binary outcome and multiple continuous or ordinal mediators: simulations and application to an alcohol intervention. Struct Equ Modeling 2016;23(3):368-383. [doi: 10.1080/10705511.2015.1062730] [Medline: 27158217]
57. Moser RP, Naveed S, Cantor D, et al. Integrative analytic methods using population-level cross-sectional data. National Cancer Institute. 2013. URL: https://hints.cancer.gov/docs/hints_ida_report.pdf [accessed 2025-07-13]
58. Hair J, Anderson R, Tatham R, Black W. Multivariate Data with Readings: Prentice Hall; 1995.
59. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
60. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. J Appl Psychol 2001 Feb;86(1):114-121. [doi: 10.1037/0021-9010.86.1.114] [Medline: 11302223]
61. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
62. Sarkar U, Karter AJ, Liu JY, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2(Suppl 2):183-196. [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]
63. Federal health IT strategic plan. Office of the National Coordinator for Health Information Technology (ONC). 2021. URL: https://www.healthit.gov [accessed 2025-07-13]
64. Charness N, Boot WR. A grand challenge for psychology: reducing the age-related digital divide. Curr Dir Psychol Sci 2022;31(2):187-193. [doi: 10.1177/09637214211068144] [Medline: 35754678]
65. Healthy people 2030. US Department of Health and Human Services. 2020. URL: https://health.gov/healthypeople [accessed 2025-07-13]
66. Czaja SJ, Boot WR, Charness N, Rogers WA. Designing for Older Adults: Principles and Creative Human Factors Approaches, 3rd edition: CRC Press; 2019. [doi: 10.1201/b22189]
67. Krist AH, Tong ST, Aycock RA, Longo DR. Engaging patients in decision-making and behavior change to promote prevention. Stud Health Technol Inform 2017;240:284-302. [doi: 10.3233/ISU-170826] [Medline: 28972524]
68. Patient empowerment strategies. Agency for Healthcare Research and Quality (AHRQ). 2019. URL: https://digital.ahrq.gov/program-overview/research-stories/engaging-empowering-patients [accessed 2025-07-13]

69. How to use HINTS data. Health Information National Trends Survey (HINTS). URL: https://hints.cancer.gov/data/download-data.aspx#H5C3 [accessed 2025-07-13]

Abbreviations

CMV: common method variance
e-commerce: electronic commerce
EMR: electronic medical record
ePHR: electronic personal health record
HINTS: Health Information National Trends Survey
HIT: health information technology
PTAM: Patient Technology Acceptance Model
SEM: structural equation model
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by H Köttl; submitted 18.01.25; peer-reviewed by G Mccormack, P Setthasuravich, S Choemprayong; revised version received 23.05.25; accepted 26.05.25; published 28.07.25.

Please cite as:
Agrawal L, DaSouza RO, Mulgund P, Chaudhary P
Frequency of Electronic Personal Health Record Use in US Older Adults: Cross-Sectional Study of a National Survey
JMIR Aging 2025;8:e71460
URL: https://aging.jmir.org/2025/1/e71460
doi:10.2196/71460

Ethics and Equity Challenges in Telerehabilitation for Older Adults: Rapid Review

Mirella Veras[1,2,3], PT, PhD; Louis-Pierre Auger[2,4], PhD; Jennifer Sigouin[2], PhD; Nahid Gheidari[5], PhD; Michelle LA Nelson[6,7], PhD; William C Miller[8,9], PhD; Anne Hudon[10], PhD; Dahlia Kairy[2,10], PhD

[1]Department of Physical Therapy, University of Manitoba, R115-771 McDermot Avenue, Winnipeg, MB, Canada
[10]Physiotherapy Program, School of Rehabilitation, Faculty of Medicine, University of Montreal, Montreal, QC, Canada
[2]The Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal, QC, Canada
[3]Centre on Aging, University of Manitoba, Winnipeg, MB, Canada
[4]School of Physical and Occupational Therapy, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada
[5]Clinical Research Unit, Montreal Neurological Institute/Hospital, Montreal, QC, Canada
[6]Lunenfeld-Tanenbaum Research Institute, Sinai Health, Toronto, ON, Canada
[7]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
[8]Occupational Science & Occupational Therapy, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
[9]Centre for Aging SMART (Solutions for Mobility, Activity, Rehabilitation and Technology), Rehabilitation Research Program, Vancouver, BC, Canada

Corresponding Author:
Mirella Veras, PT, PhD
Department of Physical Therapy, University of Manitoba, R115-771 McDermot Avenue, Winnipeg, MB, Canada

Abstract

Background: The integration of technology in rehabilitation is transforming health care delivery for older adults, especially through telerehabilitation, which addresses barriers to in-person care.

Objective: This rapid review explores the ethical and equity concerns associated with telerehabilitation for older adults, focusing on challenges such as internet access, technology adoption, and digital literacy.

Methods: Conducted according to Cochrane Rapid Review guidelines, this review used the Metaverse Equitable Rehabilitation Therapy framework, focusing on equity and ethics. Studies included telerehabilitation services for adults aged 55 years and older, published between 2010 and 2023. Screening was conducted independently by 2 researchers using Rayyan (Qatar Computing Research Institute, Hamad Bin Khalifa University), with full-text review by additional team members. Searches were performed in Medline and CINAHL

Results: From 323 papers retrieved, 49 studies met the inclusion criteria. The included studies were published between 2013 and 2023. Disparities in socioeconomic status, geographic location, and racial and ethnic backgrounds were found to impact telerehabilitation use. Additionally, ethical concerns around privacy, security, and autonomy were often inadequately addressed.

Conclusions: This review emphasizes the need for culturally appropriate, accessible, and inclusive telerehabilitation services that integrate ethical and equity considerations into their design and delivery.

(JMIR Aging 2025;8:e69660) doi:10.2196/69660

KEYWORDS

aging; older adults; equity; ethics; telerehabilitation; telehealth; review; rehabilitation; digital health; geriatric; virtual care

Introduction

The use of technology in rehabilitation is increasing worldwide, including its application in delivering services to the older adults population [1]. This is transforming how health care services are delivered to the aging population [2]. Telerehabilitation, which involves the evaluation and treatment of patients through technology, has emerged as an attractive option for older adults who may have multiple comorbidities [1,3]. Before the COVID-19 pandemic, numerous adults faced challenges accessing rehabilitation treatments due to diverse clinical limitations and geographic circumstances. Telerehabilitation has assumed greater importance in addressing the needs of older adults, as the COVID-19 pandemic exacerbated difficulties by further restricting face-to-face interactions with implementation of societal measures aimed at lowering disease prevalence [4].

When older adults experience complex medical needs, whether due to multiple chronic conditions or acute illnesses, it can lead to increased dependence on others and diminish the quality of life for both the individuals and their caregivers [5]. Health care

technologies can help overcome these challenges and enhance the well-being of the aging population by improving access to rehabilitation [5]. In this context, telerehabilitation provides an alternative platform for delivering health education and care to clients in clinical, community, or home care settings [6]. However, challenges in delivering telerehabilitation to older adults have been reported in the literature and include issues with internet access, adoption of technology, self-efficacy, experience with technology, frequency of usage, and reliance on guidance and digital literacy [7,8]. Other barriers reported in a systematic review about the use of telehealth use among older adults were trust of internet information, reading small fonts, information overload, clicking small icons, interacting with scroll bars, lack of understanding of app capabilities, fear of others overhearing, security, need for support to use, cost, privacy concerns, among others [8].

Despite the numerous benefits of telerehabilitation for older adults, this is recognized as an extremely heterogeneous population; therefore, there is a pressing need to investigate and understand the ethical and equity concerns associated with these services. No reviews to date have thoroughly explored these aspects, which is critical given the reported challenges. Knowing about these problems is important to help address the effectiveness and accessibility of telerehabilitation. Ethical concerns may arise regarding the confidentiality of patient information, while equity issues could stem from disparities in access to technology and the internet among different socioeconomic groups, for example. Addressing these concerns is essential to ensure that telerehabilitation services are delivered fairly and effectively, promoting better health outcomes and quality of life for all older adults. This review aims to examine these concerns and explore telerehabilitation for older adults over the past 10 years.

Methods

Study Design

Rapid reviews are a form of synthesis that accelerates the traditional systematic review process, facilitating the dissemination of literature in a resource-efficient manner, particularly relevant in the context of rapidly advancing technology research [9]. This rapid review aimed to explore and summarize the current literature on the equity and ethical concerns associated with delivering telerehabilitation services, following the Cochrane Rapid Review guidelines [9]. To guide the data extraction and analysis of this review, the Metaverse Equitable Rehabilitation Therapy (MERTH) framework [10] was applied, adapted to telerehabilitation, with a focus on equity (accessibility, inclusivity, diversity, fairness, and cultural relevance), health services integration (responsiveness, continuity of care, and autonomy to participate in health-related decisions), interoperability, and humanization (communication, person-centered care, and empowerment) [10]. For this article, we are focusing on equity (accessibility, inclusivity, diversity, fairness, cultural relevance) and empowerment.

We included studies that focused on telerehabilitation services for both older adults and middle-aged individuals aged 55 years and above. The inclusion criteria focused on reviews (systematic reviews, scoping reviews, meta-analyses, and narrative reviews with descriptions of included studies) involving older adults aged 55 years and older, which includes the older working-age population (55 to 64 y) [11]. This age range was selected to include individuals who may benefit from telerehabilitation services, considering their potential health care needs and technological engagement during this phase of life. These studies were selected if published between 2010 and 2023. Editorials, conference abstracts, and papers without full text available were excluded. Two researchers (MV and NG) independently completed the title and abstract screening using Rayyan [12], an online app that accelerates initial screening of abstracts and titles for managing reviews. One reviewer (MV) performed the full-text screening, and any doubts regarding inclusion were discussed with MV, DK, and JS. DK, JS, NG, and MV developed a data extraction tool and refined it throughout the extraction process. For added rigor, a sample of the data, such as 10% (13/125) of the reviews at the abstract phase, was independently reviewed by another team member (DK or JS).

To conduct the search, we consulted a specialized health sciences library. We tailored a search strategy for 2 databases: Medline and CINAHL.

Detailed search strategies can be found in Multimedia Appendix 1.

Ethical Considerations

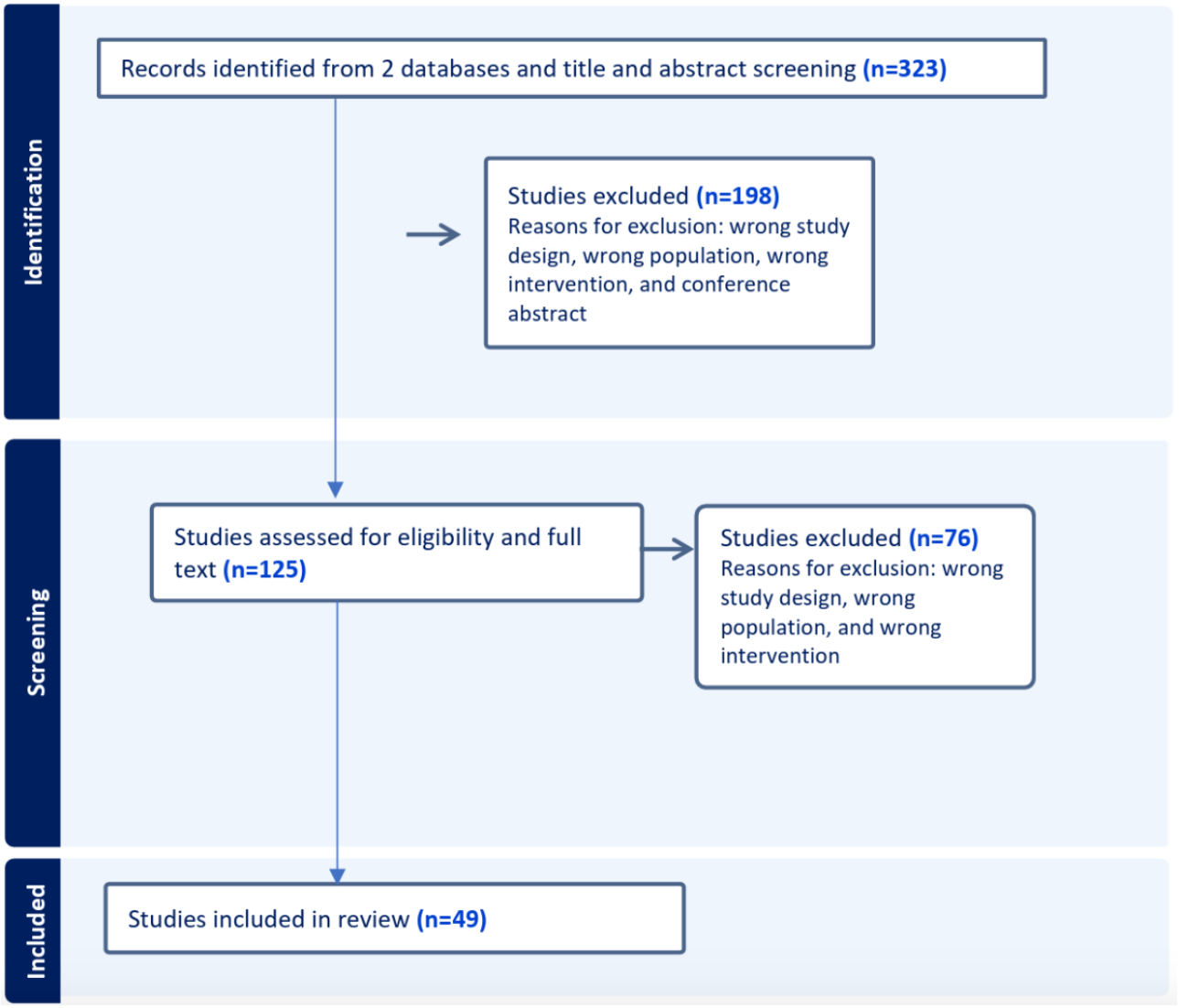
This study is a review of existing literature and does not involve any primary data collection or direct interaction with human participants. Therefore, ethical approval and consent for publication are not required. All sources cited have been appropriately referenced per academic standards.

Results

Overview

Figure 1 provides an overview of the review process. Initially, 323 papers were retrieved through the search. After title and abstract screening, followed by full-text review, 49 studies met the criteria for extraction and final evaluation.

Figure 1. Flow diagram of study selection.



Publications Characteristics

The included review studies were published between 2013 and 2023. Approximately 39% (n=19) of the studies were published in 2021. The included studies predominantly consisted of systematic reviews (22 studies), systematic reviews and meta-analyses (11 studies) followed by scoping reviews (8 studies; Table S1 in Multimedia Appendix 2 [6,13-60]).

Morbidities

Cardiorespiratory, neurological, and musculoskeletal conditions are the primary morbidities present in the studies identified in this review. Other key categories include cognitive impairment and mental health, cancer, and diabetes. Less frequently studied morbidities include vision and hearing problems, balance disorders, risk for falls, and pelvic floor disorders. The distribution of included studies (n=49) across morbidity categories revealed that the most commonly addressed condition was cardiorespiratory, featured in 33/49 (67%) of the studies. Neurological and musculoskeletal conditions followed this, each reported in 17/49 (35%) of the studies. Studies categorized under "other" comprised 15/49 (31%). Cognitive impairment and mental health, and cancer were each represented in 9/49 (19%) of the studies. Diabetes was addressed in 6/49 (13%), while vision and hearing problems, balance disorders and risk for falls, and pelvic floor disorders were the least represented, each occurring in 2/49 (5%) of the included studies. These findings indicate a predominant research focus on cardiorespiratory and neuromusculoskeletal morbidities.

It is important to note that individual morbidities can be counted in more than 1 study within these reviews and that each study could include participants with several morbidities. This repetition highlights the prevalence and importance of specific categories of health issues within the reviewed literature.

Type of Technologies Used in Telerehabilitation for Older Adults

Within the included reviews, the categorization of communication technologies in telerehabilitation included several key areas, each highlighted by different examples and applications. Communication technologies included eHealth platforms such as telehealth, telemedicine, and telerehabilitation, which used tools such as videoconferencing (Skype [Microsoft Corp] and Zoom [Zoom Communications, Inc]), telephones, text messaging, emails, and web-based platforms. Telemonitoring and remote monitoring were used through health

monitoring systems, wearable sensors, and mobile health apps for tracking health metrics. Virtual and augmented reality were applied in digital rehabilitation systems and exergames such as Wii Fit (Nintendo Entertainment Analysis & Development) and Kinect (Microsoft). Wearable technology included sensors, smartwatches, and fitness trackers. Education and support tools provided digital health education via e-learning apps and digital pain coaches. Digital and assistive technologies enhanced patient engagement and supported independent living through interactive internet platforms and emergency assistance systems. Finally, home-based rehabilitation and exercise facilitated remote rehabilitation exercises with personalized programs and activity trackers (Table 1)

Based on the studies included in this rapid review, the most frequent categories of technologies used are communication technologies, including telemonitoring and remote monitoring, which appear in 73% of the included studies. In contrast, the least used categories are virtual and augmented reality and interactive and assistive technologies, each appearing in 12% of the studies. Specifically, virtual and augmented reality were found in 6 studies, which corresponds to approximately 12% of the total studies. The total percentage exceeds 100% due to the overlap of studies across multiple categories (Table 1).

Table . Study characteristics: technologies used and examples.

Categorization	Examples	References
Communication technologies	• eHealth platforms: telehealth interventions, telemedicine, and telerehabilitation platforms • Telehealth devices: remote supervision tools, home-health monitoring systems, and teleconsultation platforms Videoconferencing: Skype (Microsoft Corp), Zoom (Zoom Communications, Inc), FaceTime (Apple Inc), Adobe Connect, and Microsoft Teams • Telephone and mobile phones: standard calls, smartphones, and telephone support • Text messaging and emails: SMS and email • Web-based platforms: online multimedia, discussion boards, blogs, and web-based health management portals	[6,13-42,60]
Telemonitoring and remote monitoring	• Health monitoring systems: heart health monitoring system, wearable sensors, pulse oximeters, and blood pressure monitors • Telemonitoring devices: ECG[a], heart monitors, Bluetooth-powered sensors, and wireless transmitters • Mobile health apps: apps for tracking health metrics and remote wound care applications	[6,13,18,21-23,26,28,30,35,37,39,43-50]
Virtual and augmented reality	• Virtual rehabilitation systems: virtual reality–based rehabilitation and interactive virtual reality exercises • Exergames: Wii Fit (Nintendo Entertainment Analysis & Development), Kinect (Microsoft), VRc[b] video game dancing, and interactive gaming systems	[6,19,23,51,52,61]
Wearable technology	• Sensors and monitors: inertial sensors, accelerometers, gyroscopes, and pedometers • Wearable devices: smartwatches, fitness trackers, and wearable pressure	[6,19,21,41,43-46,53,59]
Education and support tools	• Provides health education and support through digital tools. Examples: e-learning apps, digital pain coaches.	[6,14,15,17,21,23,26,30,36,38,45,46,50,54-56]
Interactive and assistive technologies	• Enhances patient engagement and supports independent living. Examples: interactive internet platforms, emergency assistance	[17,22,39,50,53]
Home-based rehabilitation and exercise	• Facilitates rehabilitation exercises at home with remote guidance. Examples: personalized exercise programs, activity trackers.	[6,15,21,22,24,35,43-45,47,48,50,52,53,55]

[a]ECG: electrocardiogram.

[b]VR: virtual reality.

Equity

Sex or Gender

Twenty-nine reviews included at least 1 study that reported findings on gender (Table S1 in Multimedia Appendix 2). The reviews that reported sex or gender in their results did not provide a full analysis based on these variables. However, 2 studies included some comments related to their gender or sex results.

One review reported that the findings regarding gender and the use of eHealth are mixed. “Out of 22 articles, only five suggest that gender plays a role in influencing eHealth usage, while eight studies found no connection between gender and eHealth use.” The studies did not define gender, underscoring the importance of clearly specifying sex, gender, and sexual orientation within the included population. Additionally, 3 studies highlight that women tend to be more engaged with and satisfied by eHealth apps, using them more frequently than men

[13]. In 1 review, the authors mentioned a gender bias, with more male participants than female participants [6]. In another study, the authors noted that most of the studies included more women than men. Four studies included only women. One study explained this decision as "to avoid the influence of gender differences on the risk of falling." Two studies excluded the few male participants, and another study did not provide a reason for the exclusive participation of women. Additionally, 1 study did not describe the age range or gender distribution of the participants [43].

Ethnicity

Racial differences in telehealth care use were observed, with cultural minority groups. For instance, 1 study found no significant effect of telehealth interventions on older Hispanic patients with chronic heart failure [14]. Overall, most studies that reported ethnicity included predominantly White participants. The term "Caucasian" is used as a more formal and historically established term to describe people of European descent. However, it can be viewed as outdated or less inclusive compared to simply using "White" or more specific ethnic identifiers (Table S1 in Multimedia Appendix 2).

Geographic Location

Older adults with chronic diseases residing in rural areas are less likely to use eHealth services compared to their urban counterparts, primarily due to limited access and lower socioeconomic status. Rural inhabitants face greater barriers to the effective implementation of eHealth, often related to insufficient infrastructure and resources. Contrarily, some regions, such as certain cities in South Korea with limited medical facilities, have shown a higher adoption of eHealth compared to more urbanized areas, highlighting the potential benefits of these technologies in underserved locations [13] (Table S1 in Multimedia Appendix 2).

Socioeconomic Status

One study highlighted that despite lower-income groups often facing barriers in accessing eHealth technologies due to economic constraints, their satisfaction levels with using these technologies were found to be comparable to those of higher-income groups once they had access (Reiners et al, 2019) [13]. This suggests that while income initially affects access to eHealth services, it may not impact the perceived benefits and satisfaction derived from using them once accessibility barriers are overcome.

Barriers

Digital barriers: the following summaries highlight the main challenges and associated studies for each category concerning the adoption of telerehabilitation among older adults:

Economic barriers: affordability of technology is an obstacle for older adults and those with lower educational backgrounds (Bertolazzi et al, 2024 [57] and Bhattarai and Phillips, 2017 [15]).

Usability challenges: difficulty in using mHealth technologies hinders adoption, exacerbated by lack of exposure (Johnson et al, 2021) [44].

Digital literacy disparities: lower digital literacy among racial or ethnic minorities, low-income groups, and people with sensory and manual dexterity deficits impact technology acceptance (Yi et al, 2021 [58]; Johnson et al, 2021 [44]; Jonker et al, 2020 [45]; Bertolazzi et al, 2024 [57]; Bostrom et al, 2020 [46]; Beckie, 2019 [16]; Gaspar and Lapão, 2021 [43]);

Support from family members: family encouragement aids acceptance and use of eHealth technologies (Jonker et al, 2020) [45].

Technology design and training: user-friendly designs and proper training enhance technology adoption (Johnson et al, 2021 [44] and Bhattarai and Phillips, 2017 [15]).

Inclusion and accessibility: addressing barriers such as income and education levels is crucial for equitable access (Bertolazzi et al, 2024) [57].

Health care provider involvement: clinician support and effective information flow facilitate technology adoption (Bhattarai and Phillips, 2017) [15].

Costs

The results related to digital health technology costs are structured into several key categories: costs, benefits, equity considerations, and policy implications. First, the section on cost details the financial implications of adopting digital health technologies, highlighting initial investments, ongoing maintenance costs, and cost analyses. Second, the benefits category outlines the positive impacts of these technologies on telerehabilitation, such as improved patient outcomes, enhanced efficiency, and reduced health care disparities. Third, equity considerations address how digital health technologies can potentially bridge gaps in health care access and outcomes among the older adult population. Lastly, the section on policy implications describes regulatory frameworks, privacy concerns, and recommendations for effective implementation strategies (Table S1 in Multimedia Appendix 2).

Overview of Costs in Digital Health Technology

Background of Costs in Digital Health Technology

Digital health technologies include a range of costs associated with implementation and operation. The following provides a summary of findings concerning cost considerations, supplemented with insights from the studies reviewed. For detailed examples and other comments, examples can be found in Table S1 in Multimedia Appendix 2.

General Cost Implications

Telehealth technologies are often highlighted as cost-effective strategies, particularly in reducing hospital and nursing home expenses.

Cost-Effective Strategies

For instance, telehealth has been proposed as cost-effective for delivering health education and promoting self-monitoring behaviors among older adults with chronic conditions.

Cost Types

Costs in digital health include implementation activities and operational expenses related to eHealth applications.

Specific Cost and Benefit Types

Patient Outcomes

Studies predominantly capture patient outcomes through quality-of-life considerations and physical health status indicators (summary of results of the included studies).

Resource Use

Emphasis is placed on reducing home care visits and hospital usage through remote monitoring, aimed at preventing unnecessary hospital admissions (results of included studies).

Equity and Access Considerations

Gender and Socioeconomic Variables

These variables influence the usability and adoption of digital health technologies, highlighting disparities in access and usage (comment on discussion of included studies).

Racial and Ethnic Differences

Studies note disparities in telehealth use among different racial and ethnic groups, suggesting cultural acceptance and access barriers.

Geographic and Socioeconomic Factors

Access to and effectiveness of telerehabilitation are influenced by factors such as geographic location and socioeconomic status, with rural and lower-income populations facing greater challenges.

Recommendations and Future Directions Related to Costs

Addressing Barriers

Recommendations include designing culture-friendly and accessible telehealth services to enhance engagement among diverse populations.

Policy Implications

Policy makers are urged to minimize cost barriers, provide subsidies for technology access, and promote equitable health technology adoption.

Ethics

Overview

Twenty-five studies reported some ethical aspects of the included studies. The following ethical aspects of the included studies can be categorized into safety, adverse events, privacy, empowerment, respect and knowledge of cultural diversity, and inclusion of ethics in research design.

Safety and Adverse Events

Safety concerns focus on ensuring that interventions do not cause harm to participants. Studies highlight that most interventions were safe, with no substantial adverse events reported (Huang et al, 2020 [17]; Malaguti et al, 2021 [18]; Johnson et al, 2021 [44]; Kraaijkamp et al, 2021 [19]; Bostrom et al, 2020 [46]; Gaspar and Lapão, 2021 [43]; Reeder et al, 2016 [20]; Solis-Navarro et al, 2022 [31]). The adverse events category pertains to the reporting and management of negative outcomes associated with interventions. Some studies mentioned the importance of feasibility testing to identify potential adverse events (Dennette et al, 2021 [21]; Kraaijkamp et al, 2021 [19]; Devi et al, 2015 [22]; Dequanter et al, 2021 [23]; Ambrens et al, 2022 [24]; Del Pino et al, 2022 [47]; and Ding et al, 2023 [25]).

Privacy

Concerns about maintaining the confidentiality and privacy of participant data were noted, emphasizing the need for secure data handling and adherence to privacy regulations (Yi et al, 2021 [58] and Jonker et al, 2020 [45]).

Respect and Knowledge of Cultural Diversity

This category highlights the importance of culturally appropriate interventions, especially for Indigenous older adults, and includes the use of culturally relevant content and rewards (Choukou et al, 2021 [26]).

Empowerment

In 1 study, the authors highlighted the empowerment divide, noting that despite having access to the necessary hardware and skills to use eHealth apps, some individuals did not use these opportunities because they did not feel personally empowered or believe they would benefit. While the reviewed papers primarily addressed economic and usability divides, the authors emphasized that future research should specifically focus on overcoming the empowerment divide (Reiners et al, 2019 [13]).

Inclusion of Ethics in the Design of the Interventions

Several studies highlighted in their discussion the importance of incorporating ethical considerations in the design and implementation of interventions. This includes respecting client preferences, obtaining informed consent, and addressing ethical and privacy concerns (Saito and Izawa, 2021 [27]; Sulz et al, 2021 [28], Su et al, 2020 [29]; Theodoros et al, 2019 [30]; and Tao et al, 2018 [42]).

Our review showed that 26 studies did not address the ethical aspects of their interventions. However, among those that did, several key ethical considerations emerged. Safety was a primary focus, with studies generally reporting no adverse events associated with the interventions (Huang et al, 2020 [17]; Malaguti et al, 2021 [18]; Johnson et al, 2021 [44]; Kraaijkamp et al, 2021 [19]; Bostrom et al, 2020 [46]; Gaspar and Lapão, 2021 [43]; Reeder et al, 2016 [20]; Solis-Navarro et al, 2022 [31]). Adverse event management was emphasized through feasibility testing (Dennette et al, 2021 [21]; Kraaijkamp et al, 2021 [19]; Devi et al, 2015 [22]; Dequanter et al, 2021 [23]; Ambrens et al, 2022 [24]; Del Pino et al, 2022 [47]; Ding et al, 2023 [25]).

Privacy concerns were frequently mentioned, highlighting the need for secure data handling and compliance with privacy regulations (Yi et al, 2021 [58]; Jonker et al, 2020 [45]). Empowerment was also critical, with studies stressing the importance of making participants feel capable and confident in using eHealth technologies (Reiners et al, 2019 [13]). Respect

and knowledge of cultural diversity were crucial, particularly for indigenous older adults, calling for culturally appropriate content and rewards (Choukou et al, 2021) [26]. Finally, the inclusion of ethics in research design was recommended, advocating for respect for client preferences, informed consent, and addressing ethical and privacy issues throughout the intervention process (Saito and Izawa, 2021 [27]; Sulz et al, 2021 [28]; Su Jing, 2020 [29]; Theodoros et al, 2019 [30]; and Tao et al, 2018 [42]). These findings emphasize the necessity of integrating comprehensive ethical considerations into digital health interventions to ensure safety, privacy, and respect for all participants.

Discussion

Principal Findings

The distribution of morbidities highlights areas where research is concentrated and indicates potential gaps in the literature that warrant further investigation, such as vision and hearing problems, balance disorders, risk of falls, and pelvic floor disorders. This distribution might also reflect conditions that health professionals are more comfortable managing. Therefore, there is a need to develop strategies to offer remote rehabilitation services for these conditions as well.

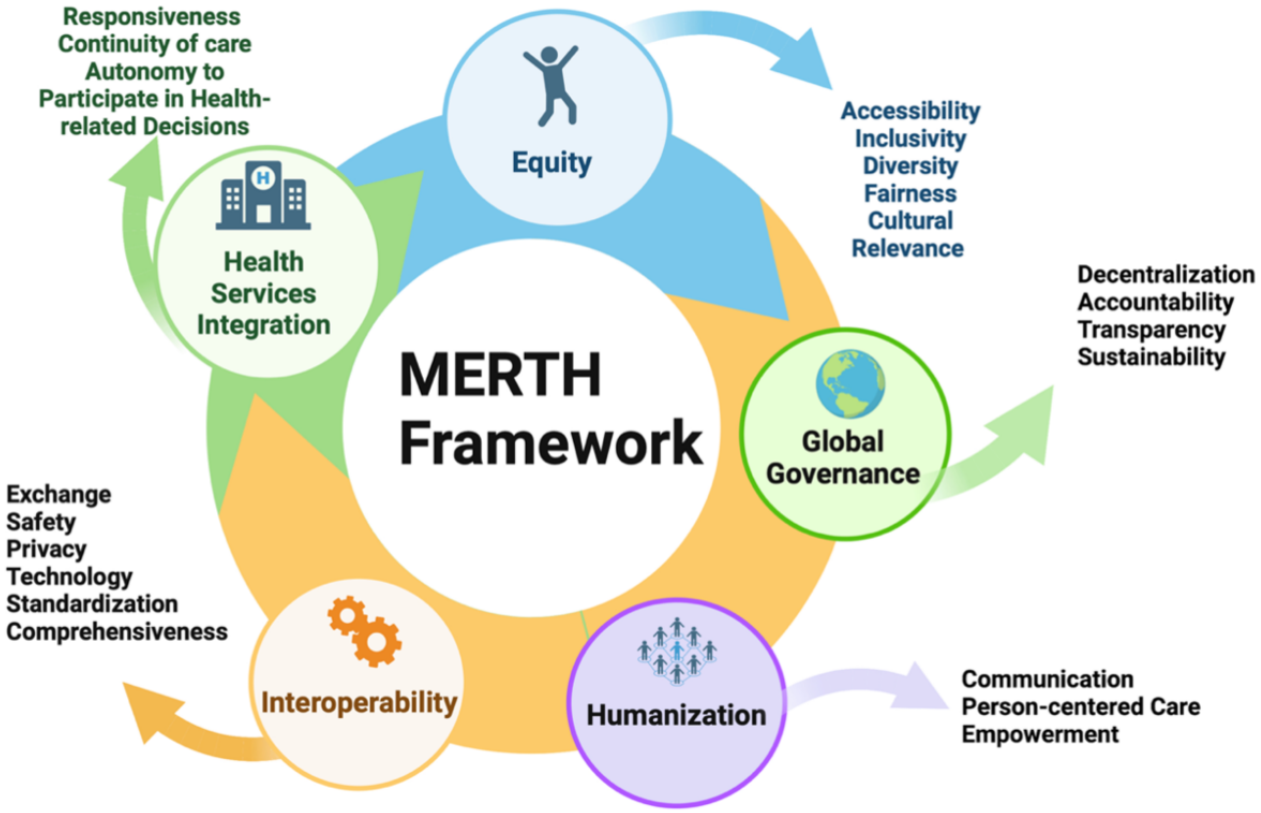
Older adults have a higher prevalence of chronic obstructive pulmonary disease, a trend that is expected to rise substantially in the coming decades due to aging populations and prolonged exposure to risk factors [62]. It has been estimated that between 55% and 98% of older adults aged 60 years and older have at least 2 chronic diseases, known as multimorbidity. Among these, cardiovascular diseases are the most common [63]. This aligns with our findings that cardiorespiratory issues were the most frequently studied morbidity type.

When considering the Health Services Integration of the MERTH framework (Figure 2), the need for continuity of care and responsiveness becomes even more evident [10]. To effectively address the complex health needs of older adults, particularly those with multiple chronic conditions, health systems must be able to provide integrated and coordinated care across various settings and providers. This not only involves traditional clinical care but also includes social support systems and digital resources, which are increasingly important for managing aging-related health issues [10].

Incorporating telerehabilitation could be a key part of achieving this integration. These digital platforms can help bridge gaps in health care access and ensure that patients receive consistent care, no matter their location or mobility [10]. They also offer an innovative way to enhance autonomy in decision-making, allowing patients to engage in their rehabilitation process more actively and with greater flexibility. This is especially important as older adults may have difficulties attending in-person sessions due to transportation issues or other physical limitations.

In the rapid review, communication technologies, telemonitoring, and remote monitoring were the most frequently included categories, totaling 38 reviews. In contrast, virtual and augmented reality, as well as interactive and assistive technologies, were less frequently cited, with only 6 reviews each. This distribution indicates a stronger focus on technologies that enable direct communication and continuous health monitoring, while immersive and interactive technologies, though valuable, are less commonly studied.

Figure 2. MERTH framework [10]. MERTH: Metaverse Equitable Rehabilitation Therapy.



This review highlights several key factors influencing the use and effectiveness of telerehabilitation services across different demographics. Racial disparities are evident, with cultural minority groups often showing lower acceptance of telehealth. For instance, 1 study found no significant impact of telehealth interventions on older Hispanic patients with chronic heart failure, indicating that cultural differences may influence outcomes. There is a substantial research gap in addressing racial differences in telerehabilitation, and the current literature primarily identifies barriers without proposing solutions to promote racial equity in telerehabilitation use. To ensure successful interventions, health care providers must design and deliver culturally friendly telerehabilitation services tailored to their target patients. Future research should focus on exploring these racial differences further and developing strategies to enhance the acceptance and effectiveness of telerehabilitation among diverse racial and ethnic groups.

Geographic location also plays a crucial role, with older adults in rural areas less likely to use telerehabilitation services due to limited access and lower socioeconomic status. These rural barriers often stem from insufficient infrastructure and resources, although some regions, such as certain South Korean cities with limited medical facilities, have shown higher eHealth adoption, demonstrating the potential of these technologies in underserved areas [13]. The overall trend emphasizes a persistent urban-rural divide in telerehabilitation usage, necessitating targeted efforts to improve access and infrastructure in rural areas to ensure equitable health care delivery.

Socioeconomic status further influences telerehabilitation use; lower-income groups face initial access barriers but report satisfaction levels comparable to higher-income groups once they gain access, indicating that economic constraints primarily affect initial accessibility rather than ongoing satisfaction and perceived benefits. The study on gender disparities in telehealth use among older adults in the United States during the COVID-19 pandemic provides relevant results for comparison with our review. It showed a significant increase in telehealth use among females during the pandemic, influenced by factors such as multimorbidity, tablet ownership, and learning new technologies. Males with similar factors also had higher odds of telehealth use [64]. Although some reviews in our study mentioned gender differences, they did not provide much detailed information. These findings emphasize the necessity for targeted interventions to enhance older adults' access to telerehabilitation services and mitigate digital disparities, aligning with our review's emphasis on the need for equitable access across diverse populations.

The reviews included in this analysis did not provide a detailed discussion on gender, sexual orientation, or the distinction between sex and gender. While the term "gender" may have been used in some instances, it was not consistently defined or explored in the context of the studies reviewed. In fact, many of the reviews primarily focused on sex-based differences, and in some cases, the term "gender" may have been used interchangeably with sex without clarifying the distinction. Additionally, there was no substantial examination of sexual orientation, and these factors were not adequately addressed or analyzed across the included studies. To provide a more comprehensive understanding of how these variables influence telerehabilitation outcomes, future research should ensure a clear and explicit discussion of gender, sex, and sexual

orientation, and their potential impact on the accessibility and effectiveness of telerehabilitation services.

Digital literacy plays a significant role in the adoption of telerehabilitation among older adults, with several key challenges highlighted in this review. Economic barriers, such as the affordability of technology, pose significant obstacles, particularly for older adults and those with lower educational backgrounds. Usability challenges, often exacerbated by a lack of exposure to health technologies, further hinder adoption. Studies show that enhancing internet use among older adults can significantly improve their access to health information and their ability to manage their health [65,66].

Disparities in digital literacy are particularly pronounced among racial or ethnic minorities, low-income groups, and individuals with sensory and manual dexterity deficits, affecting technology acceptance and use. Family support has been shown to aid in the acceptance and use of health technologies, while user-friendly design and proper training can significantly enhance adoption rates. A randomized controlled trial on technophilia (enthusiasm toward new technologies) among people with dementia or mild cognitive impairment and their caregivers demonstrated technophilia was lower among persons with disabilities influenced by factors such as younger age, male gender, higher education, better health, and depression. For caregivers, lower burden and better quality of life also contributed [67]. The sensory and manual dexterity deficits can significantly impact how individuals accept and use technology. These challenges highlight the importance of designing inclusive technology solutions that accommodate a diverse range of physical abilities, ensuring equitable access and usability for all users. Addressing all these barriers is crucial for ensuring equitable access to these technologies.

In terms of accountability, as per the MERTH framework, telerehabilitation services must maintain high standards of care despite these barriers. One area where accountability becomes particularly important is in ensuring that health care providers are properly trained and equipped to offer culturally competent and accessible digital care [10]. As there is a disparity in the acceptance and effectiveness of telerehabilitation across different demographics—such as rural versus urban populations, and racial or ethnic minorities—health care providers must be held accountable for creating inclusive services. This includes adapting technologies to be more accessible, as well as implementing feedback mechanisms to assess the effectiveness of interventions.

Cost considerations include initial investments and ongoing maintenance, but digital health technologies are generally seen as cost-effective, particularly in reducing hospital and nursing home expenses [14,68,69]. To maximize benefits and ensure equity, it is important to design culturally friendly and accessible telerehabilitation services, minimize cost barriers, and promote equitable technology adoption across diverse populations [14]. These findings support our review results on the need for targeted interventions to improve access to telerehabilitation services and reduce digital disparities among older adults.

Regarding ethical aspects, 26 studies did not address the ethical aspects of their interventions. This gap in addressing ethical considerations highlights a substantial shortcoming in the research, underscoring the need for more rigorous ethical evaluation in future studies. However, for the studies that did address ethics, several key considerations came to light. Safety was reported as a comment or result with no significant adverse events associated with the interventions. Adverse event management was emphasized through feasibility testing. A recent scoping review of risks, adverse effects, and mitigation strategies in delivering mental health services via telehealth, emphasizes the growing recognition of the importance of adverse event reporting in telehealth. However, it also highlights the lack of standardized protocols for this purpose [61]. Challenges persist in the telerehabilitation regarding the systematic collection and evaluation of adverse events. The study urges improved reporting of near-misses and actual incidents to enhance risk management. It also advocates for comprehensive clinical training to prevent adverse events and proposes the establishment of robust reporting mechanisms to promote ongoing enhancements in telehealth practices [61]. There is a clear and pressing need to establish standardized protocols for reporting adverse events in telerehabilitation, as highlighted by the findings. This will enable better risk management and improve the overall quality and safety of telerehabilitation services.

The review authors emphasized the importance of addressing privacy concerns and noted a lack of attention to these issues in the included studies. This highlights the need for secure data handling and adherence to privacy regulations. The studies reviewed here briefly addressed privacy and security concerns but did not thoroughly investigate these aspects. Privacy risks in telerehabilitation practice include environmental factors (such as inadequate private spaces for sensitive consultations), technology issues (including data security and limited internet access), and operational challenges [70]. The privacy concerns raised in the review align closely with elements of the MERTH framework: interoperability, particularly regarding exchange, safety, privacy, technology standardization, and comprehensiveness [10]. The challenges highlighted, such as inadequate private spaces, data security issues, and limited internet access are foundational to the safe and effective exchange of information in telerehabilitation. These issues underscore the importance of prioritizing robust data handling and adherence to privacy regulations to safeguard sensitive information of older adults [10].

Empowerment was also critical, with studies stressing the importance of making participants feel capable and confident in using health technologies. There is a noticeable gap between the digital technologies developed for older adults and what they actually need. Aging is often viewed negatively, with older adults stereotypically seen as frail and incompetent [20]. Consequently, many technologies designed for this demographic focus primarily on care rather than empowerment. These negative stereotypes commonly result in excluding older adults from the research and design processes of digital technology. Mannheim et al [71] argue that including older adults in these processes is crucial for creating technologies that genuinely enhance their well-being. Age limits are frequently justified by ageist assumptions that question the competence and reliability

of older adults. Social exclusion, driven by these stereotypes, is evident in various areas, including access to services, social relations, and research participation. Arbitrary upper age limits, which more than half of the relevant studies impose, often result in the unjustified exclusion of older adults from clinical research and randomized controlled trials. Although the use of these exclusion criteria may be declining, selective exclusion remains common [71]. These practices highlight ethical challenges in recognizing older adults' ability to participate meaningfully in research and design, underscoring the impact of stereotypes on their involvement.

This issue aligns closely with the MERTH framework's element humanization, which emphasizes the importance of person-centered care, communication, and empowerment in health care [10]. The humanization of health care, especially in the context of telerehabilitation, means creating environments where patients are treated as whole individuals, not just as a collection of symptoms or conditions. This approach fosters respect, dignity, and empowerment, encouraging patients to engage actively in their own care and decision-making. In the case of telerehabilitation and digital health technologies, this translates to providing digital spaces that not only cater to physical rehabilitation but also consider emotional, social, and psychological needs [10].

Respect and knowledge of cultural diversity were crucial, particularly for Indigenous older adults, calling for culturally appropriate content and rewards [26]. There is a lack of studies that include Indigenous older adult populations. A study conducted with Indigenous older adults in Saskatchewan showed that many express a desire to learn more about technology and recognize its value in supporting healthy aging [26]. The study highlighted that some researchers emphasized the importance of Indigenous language apps and the revitalization of cultural teachings, seeking technology's assistance in facilitating these cultural learning processes. It emphasizes the importance of considering the role of families in engaging older adults with technology, given the crucial cultural role families play within Indigenous communities [26].

The findings of this review underline the necessity of integrating comprehensive ethical considerations into telerehabilitation interventions to ensure safety, privacy, and respect for all populations.

Study Strengths and Limitations

The strengths of this rapid review include its comprehensive synthesis of a wide range of literature within a condensed timeframe, providing timely insights into telerehabilitation for older adults. This efficiency allows for a quick overview of key findings and trends, aiding health care practitioners and policy makers in making informed decisions promptly. Additionally, the review's focus on ethics and equity in telerehabilitation ensures that critical considerations around patient safety, privacy, and inclusivity are highlighted, promoting ethical standards in telerehabilitation delivery. Finally, by emphasizing the integration of diverse perspectives and the identification of gaps in current research, this rapid review contributes to advancing knowledge and guiding future investigations aimed at enhancing telerehabilitation effectiveness and accessibility for older adults.

Study limitations include language constraints, as searches were limited to English and French publications. Future research should explore strategies to access studies in all languages, particularly in equity contexts, to ensure broader inclusivity. Additionally, the review did not ascertain the specific rehabilitation professionals involved in the studies, as individual study details were not accessed, and this information was, for the most part, not presented in the reviews. Furthermore, data extraction was conducted by a single reviewer, which may have introduced potential biases. To address this, the single reviewer discussed the results with the team and engaged in iterative discussions to mitigate any potential biases. Additionally, to further ensure accuracy, 10% of the reviews, including abstracts, were independently cross-checked by another team member. These limitations highlight the need for comprehensive approaches in future reviews to enhance inclusivity and rigor in synthesizing evidence across diverse contexts and disciplines.

Conclusions

In conclusion, this review has shed light on the distribution of studies across various health conditions within telerehabilitation contexts, emphasizing a predominant focus on cardiorespiratory, neurological, and musculoskeletal conditions. These findings highlight both the areas of concentrated research and gaps, particularly in less-studied morbidities such as vision and hearing problems, balance disorders, and pelvic floor disorders.

The review also emphasizes disparities in technology adoption and research participation among different demographic groups among older and middle-aged adults, particularly highlighting racial disparities and socioeconomic influences on telerehabilitation use. Challenges persist in addressing these disparities, with rural areas facing barriers due to limited infrastructure and resources. Moreover, the review has identified ethical considerations as paramount, emphasizing the need for standardized protocols in adverse event reporting, improved privacy, and inclusive research practices that recognize and respect cultural diversity, including Indigenous older adults. Moving forward, addressing these gaps and challenges is essential to advancing equitable access to effective telerehabilitation services for older adults across diverse populations.

Acknowledgments

We thank the Canadian Institutes for Health Research, Association Médicale Canadienne, and Réseau Québécois de Recherche en Adaptation-Réadaptation (REPAR).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

MV wrote the first and final version of this paper. MV, JS, NG, AH, and DK were involved in the methodology, data screening, extraction, and analysis of the results. All authors (MV, JS, NG, LPA, MLAN, WCM, AH, and DK) contributed critically to all phases of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for this paper.
[DOCX File, 130 KB - aging_v8i1e69660_app1.docx]

Multimedia Appendix 2
Study characteristics of included studies.
[DOCX File, 129 KB - aging_v8i1e69660_app2.docx]

References

1. Marzaleh MA, Peyravi M, Azhdari N, Bahaadinbeigy K, Sarpourian F. Application of telerehabilitation for older adults during the COVID-19 pandemic: a systematic review. Disaster Med Public Health Prep 2023;17:e402. [doi: 10.1017/dmp.2022.219] [Medline: 36004522]
2. Roots RK, Brown H, Bainbridge L, Li LC. Rural rehabilitation practice: perspectives of occupational therapists and physical therapists in British Columbia, Canada. Rural Remote Health 2014;14:2506. [Medline: 24528153]
3. Pol M, Qadeer A, van Hartingsveldt M, Choukou MA. Perspectives of rehabilitation professionals on implementing a validated home telerehabilitation intervention for older adults in geriatric rehabilitation: multisite focus group study. JMIR Rehabil Assist Technol 2023 Jul 18;10:e44498. [doi: 10.2196/44498] [Medline: 37463040]
4. Omboni S, McManus RJ, Bosworth HB, et al. Evidence and recommendations on the use of telemedicine for the management of arterial hypertension: an international expert position paper. Hypertension 2020 Nov;76(5):1368-1383. [doi: 10.1161/HYPERTENSIONAHA.120.15873] [Medline: 32921195]
5. Gokalp H, de Folter J, Verma V, Fursse J, Jones R, Clarke M. Integrated telehealth and telecare for monitoring frail elderly with chronic disease. Telemed J E Health 2018 Dec;24(12):940-957. [doi: 10.1089/tmj.2017.0322] [Medline: 30129884]
6. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. PLoS One 2019;14(11):e0225150. [doi: 10.1371/journal.pone.0225150] [Medline: 31714924]
7. Md Fadzil NH, Shahar S, Rajikan R, et al. A scoping review for usage of telerehabilitation among older adults with mild cognitive impairment or cognitive frailty. Int J Environ Res Public Health 2022 Mar 28;19(7):4000. [doi: 10.3390/ijerph19074000] [Medline: 35409683]
8. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020 Aug 12;8(8):e20359. [doi: 10.2196/20359] [Medline: 32784177]
9. Garritty C, Gartlehner G, Nussbaumer-Streit B, et al. Cochrane Rapid Reviews Methods Group offers evidence-informed guidance to conduct rapid reviews. J Clin Epidemiol 2021 Feb;130:13-22. [doi: 10.1016/j.jclinepi.2020.10.007] [Medline: 33068715]
10. Veras M, Labbé DR, Furlano J, et al. A framework for equitable virtual rehabilitation in the metaverse era: challenges and opportunities. Front Rehabil Sci 2023;4:1241020. [doi: 10.3389/fresc.2023.1241020] [Medline: 37691912]
11. Kim YM, Jang SN, Cho SI. Working hours, social engagement, and depressive symptoms: an extended work-life balance for older adults. BMC Public Health 2024;23(1):23. [doi: 10.1186/s12889-023-17072-x]
12. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
13. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of ehealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
14. Guo Y, Albright D. The effectiveness of telehealth on self-management for older adults with a chronic condition: a comprehensive narrative review of the literature. J Telemed Telecare 2018 Jul;24(6):392-403. [doi: 10.1177/1357633X17706285] [Medline: 28449619]

15. Bhattarai P, Phillips JL. The role of digital health technologies in management of pain in older people: an integrative review. Arch Gerontol Geriatr 2017;68:14-24. [doi: 10.1016/j.archger.2016.08.008] [Medline: 27584871]
16. Beckie TM. Utility of home-based cardiac rehabilitation for older adults. Clin Geriatr Med 2019 Nov;35(4):499-516. [doi: 10.1016/j.cger.2019.07.003] [Medline: 31543181]
17. Huang J, Han Y, Wei J, et al. The effectiveness of the Internet-based self-management program for cancer-related fatigue patients: a systematic review and meta-analysis. Clin Rehabil 2020 Mar;34(3):287-298. [doi: 10.1177/0269215519889394] [Medline: 31793340]
18. Malaguti C, Dal Corso S, Janjua S, Holland AE. Supervised maintenance programmes following pulmonary rehabilitation compared to usual care for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2021 Aug 17;8(8):CD013569. [doi: 10.1002/14651858.CD013569.pub2] [Medline: 34404111]
19. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
20. Reeder B, Chung J, Stevens-Lapsley J. Current telerehabilitation research with older adults at home: an integrative review. J Gerontol Nurs 2016 Oct 1;42(10):15-20. [doi: 10.3928/00989134-20160201-02] [Medline: 26870988]
21. Dennett EJ, Janjua S, Stovold E, Harrison SL, McDonnell MJ, Holland AE. Tailored or adapted interventions for adults with chronic obstructive pulmonary disease and at least one other long-term condition: a mixed methods review. Cochrane Database Syst Rev 2021 Jul 26;7(7):CD013384. [doi: 10.1002/14651858.CD013384.pub2] [Medline: 34309831]
22. Devi R, Singh SJ, Powell J, Fulton EA, Igbinedion E, Rees K. Internet-based interventions for the secondary prevention of coronary heart disease. Cochrane Database Syst Rev 2015 Dec 22;2015(12):CD009386. [doi: 10.1002/14651858.CD009386.pub2] [Medline: 26691216]
23. Dequanter S, Gagnon MP, Ndiaye MA, et al. The effectiveness of e-Health solutions for aging with cognitive impairment: a systematic review. Gerontologist 2021 Sep 13;61(7):e373-e394. [doi: 10.1093/geront/gnaa065] [Medline: 32525977]
24. Ambrens M, Alley S, Oliveira JS, et al. Effect of eHealth-delivered exercise programmes on balance in people aged 65 years and over living in the community: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2022 Jun 10;12(6):e051377. [doi: 10.1136/bmjopen-2021-051377] [Medline: 35688586]
25. Ding J, Yang Y, Wu X, Xiao B, Ma L, Xu Y. The telehealth program of occupational therapy among older people: an up-to-date scoping review. Aging Clin Exp Res 2023 Jan;35(1):23-40. [doi: 10.1007/s40520-022-02291-w] [Medline: 36344805]
26. Choukou MA, Maddahi A, Polyvyana A, Monnin C. Digital health technology for Indigenous older adults: a scoping review. Int J Med Inform 2021 Apr;148:104408. [doi: 10.1016/j.ijmedinf.2021.104408] [Medline: 33609927]
27. Saito T, Izawa KP. Effectiveness and feasibility of home-based telerehabilitation for community-dwelling elderly people in Southeast Asian countries and regions: a systematic review. Aging Clin Exp Res 2021 Oct;33(10):2657-2669. [doi: 10.1007/s40520-021-01820-3] [Medline: 33765258]
28. Sülz S, van Elten HJ, Askari M, Weggelaar-Jansen AM, Huijsman R. eHealth applications to support independent living of older persons: scoping review of costs and benefits identified in economic evaluations. J Med Internet Res 2021 Mar 9;23(3):e24363. [doi: 10.2196/24363] [Medline: 33687335]
29. Su JJ, Yu DSF, Paguio JT. Effect of eHealth cardiac rehabilitation on health outcomes of coronary heart disease patients: a systematic review and meta-analysis. J Adv Nurs 2020 Mar;76(3):754-772. [doi: 10.1111/jan.14272] [Medline: 31769527]
30. Theodoros D, Aldridge D, Hill AJ, Russell T. Technology-enabled management of communication and swallowing disorders in Parkinson's disease: a systematic scoping review. Int J Lang Commun Disord 2019 Mar;54(2):170-188. [doi: 10.1111/1460-6984.12400] [Medline: 29923267]
31. Solis-Navarro L, Gismero A, Fernández-Jané C, et al. Effectiveness of home-based exercise delivered by digital health in older adults: a systematic review and meta-analysis. Age Ageing 2022 Nov 2;51(11):11. [doi: 10.1093/ageing/afac243] [Medline: 36346736]
32. Tschoepe R, Benfield A, Posey R, Mercer V. A systematic review of the effects of community transition programs on quality of life and hospital readmissions for adults with traumatic spinal cord injury. Arch Phys Med Rehabil 2022 May;103(5):1013-1022. [doi: 10.1016/j.apmr.2021.08.002] [Medline: 34464608]
33. van Egmond MA, van der Schaaf M, Vredeveld T, et al. Effectiveness of physiotherapy with telerehabilitation in surgical patients: a systematic review and meta-analysis. Physiotherapy 2018 Sep;104(3):277-298. [doi: 10.1016/j.physio.2018.04.004] [Medline: 30030037]
34. Whitfield T, Barnhofer T, Acabchuk R, et al. The effect of mindfulness-based programs on cognitive function in adults: a systematic review and meta-analysis. Neuropsychol Rev 2022 Sep;32(3):677-702. [doi: 10.1007/s11065-021-09519-y] [Medline: 34350544]
35. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. Patient Educ Couns 2013 Apr;91(1):14-24. [doi: 10.1016/j.pec.2012.10.018] [Medline: 23194823]

36. Curry J, Patterson M, Greenley S, Pearson M, Forbes CC. Feasibility, acceptability, and efficacy of online supportive care for individuals living with and beyond lung cancer: a systematic review. Support Care Cancer 2021 Nov;29(11):6995-7011. [doi: 10.1007/s00520-021-06274-x] [Medline: 34008080]
37. Brick R, Turner R, Bender C, et al. Impact of non-pharmacological interventions on activity limitations and participation restrictions in older breast cancer survivors: a scoping review. J Geriatr Oncol 2022 Mar;13(2):132-142. [doi: 10.1016/j.jgo.2021.09.010] [Medline: 34598902]
38. Guan T, Qan'ir Y, Song L. Systematic review of illness uncertainty management interventions for cancer patients and their family caregivers. Support Care Cancer 2021 Aug;29(8):4623-4640. [doi: 10.1007/s00520-020-05931-x] [Medline: 33495851]
39. Chan JKY, Klainin-Yobas P, Chi Y, Gan JKE, Chow G, Wu XV. The effectiveness of e-interventions on fall, neuromuscular functions and quality of life in community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Stud 2021 Jan;113:103784. [doi: 10.1016/j.ijnurstu.2020.103784] [Medline: 33120138]
40. Bacanoiu MV, Danoiu M. New strategies to improve the quality of life for normal aging versus pathological aging. J Clin Med 2022 Jul 20;11(14):14. [doi: 10.3390/jcm11144207] [Medline: 35887969]
41. Zhang J, Yang M, Ge Y, Ivers R, Webster R, Tian M. The role of digital health for post-surgery care of older patients with hip fracture: a scoping review. Int J Med Inform 2022 Apr;160:104709. [doi: 10.1016/j.ijmedinf.2022.104709] [Medline: 35144101]
42. Tao KFM, Brennan-Jones CG, Capobianco-Fava DM, et al. Teleaudiology services for rehabilitation with hearing aids in adults: a systematic review. J Speech Lang Hear Res 2018 Jul 13;61(7):1831-1849. [doi: 10.1044/2018_JSLHR-H-16-0397] [Medline: 29946688]
43. Gaspar AGM, Lapão LV. eHealth for addressing balance disorders in the elderly: systematic review. J Med Internet Res 2021;23(4):e22215. [doi: 10.2196/22215]
44. Johnson AJ, Palit S, Terry EL, et al. Managing osteoarthritis pain with smart technology: a narrative review. Rheumatol Adv Pract 2021;5(1):rkab021. [doi: 10.1093/rap/rkab021] [Medline: 33928214]
45. Jonker LT, Haveman ME, de Bock GH, van Leeuwen BL, Lahr MMH. Feasibility of perioperative eHealth interventions for older surgical patients: a systematic review. J Am Med Dir Assoc 2020 Dec;21(12):1844-1851. [doi: 10.1016/j.jamda.2020.05.035] [Medline: 32694000]
46. Bostrom J, Sweeney G, Whiteson J, Dodson JA. Mobile health and cardiac rehabilitation in older adults. Clin Cardiol 2020 Feb;43(2):118-126. [doi: 10.1002/clc.23306] [Medline: 31825132]
47. Del Pino R, Díez-Cirarda M, Ustarroz-Aguirre I, et al. Costs and effects of telerehabilitation in neurological and cardiological diseases: a systematic review. Front Med (Lausanne) 2022;9:832229. [doi: 10.3389/fmed.2022.832229] [Medline: 36523783]
48. Linn N, Goetzinger C, Regnaux JP, et al. Digital health interventions among people living with frailty: a scoping review. J Am Med Dir Assoc 2021 Sep;22(9):1802-1812. [doi: 10.1016/j.jamda.2021.04.012] [Medline: 34000266]
49. Marzano G, Lubkina V. Usability in social telerehabilitation systems for elderly users. Public Health 2017 Mar;144:1-3. [doi: 10.1016/j.puhe.2016.11.011] [Medline: 28274368]
50. Amirova A, Fteropoulli T, Williams P, Haddad M. Efficacy of interventions to increase physical activity for people with heart failure: a meta-analysis. Open Heart 2021 Jun;8(1):e001687. [doi: 10.1136/openhrt-2021-001687] [Medline: 34108272]
51. Velayati F, Ayatollahi H, Hemmat M. A systematic review of the effectiveness of telerehabilitation interventions for therapeutic purposes in the elderly. Methods Inf Med 2020 May;59(2-03):104-109. [doi: 10.1055/s-0040-1713398] [Medline: 32629502]
52. Emmanouilidis S, Hackney ME, Slade SC, Heng H, Jazayeri D, Morris ME. Dance is an accessible physical activity for people with Parkinson's disease. Parkinsons Dis 2021;2021:7516504. [doi: 10.1155/2021/7516504] [Medline: 34721836]
53. Ollevier A, Aguiar G, Palomino M, Simpelaere IS. How can technology support ageing in place in healthy older adults? A systematic review. Public Health Rev 2020 Nov 23;41(1):26. [doi: 10.1186/s40985-020-00143-4] [Medline: 33292707]
54. Allida S, Du H, Xu X, et al. mHealth education interventions in heart failure. Cochrane Database Syst Rev 2020 Jul 2;7(7):CD011845. [doi: 10.1002/14651858.CD011845.pub2] [Medline: 32613635]
55. Xie SH, Wang Q, Wang LQ, Wang L, Song KP, He CQ. Effect of internet-based rehabilitation programs on improvement of pain and physical function in patients with knee osteoarthritis: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Jan 5;23(1):e21542. [doi: 10.2196/21542] [Medline: 33399542]
56. Alaimo C, Campana E, Stoppelli MR, et al. Cognitive tele-enhancement in healthy older adults and subjects with subjective memory complaints: a review. Front Neurol 2021;12:650553. [doi: 10.3389/fneur.2021.650553] [Medline: 34290660]
57. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
58. Yi JS, Pittman CA, Price CL, Nieman CL, Oh ES. Telemedicine and dementia care: a systematic review of barriers and facilitators. J Am Med Dir Assoc 2021 Jul;22(7):1396-1402. [doi: 10.1016/j.jamda.2021.03.015] [Medline: 33887231]
59. Johnsen TL, Johansen T, Momsen AMH, et al. eHealth interventions to facilitate work participation: a scoping review. JBI Evid Synth 2021 Oct;19(10):2739-2759. [doi: 10.11124/JBISRIR-D-19-00433] [Medline: 34224522]

60. Antunes TPC, de Oliveira ASB, Hudec R, et al. Assistive technology for communication of older adults: a systematic review. Aging Ment Health 2019 Apr;23(4):417-427. [doi: 10.1080/13607863.2018.1426718] [Medline: 29451399]
61. Martiniuk A, Toepfer A, Lane-Brown A. A review of risks, adverse effects and mitigation strategies when delivering mental health services using telehealth. J Ment Health 2024 Jun;33(3):415-438. [doi: 10.1080/09638237.2023.2182422] [Medline: 36866784]
62. Matera MG, Hanania NA, Maniscalco M, Cazzola M. Pharmacotherapies in Older Adults with COPD: Challenges and Opportunities. Drugs Aging 2023 Jul;40(7):605-619. [doi: 10.1007/s40266-023-01038-0] [Medline: 37316689]
63. Marengoni A, Angleman S, Meinow B, et al. Coexisting chronic conditions in the older population: variation by health indicators. Eur J Intern Med 2016 Jun;31:29-34. [doi: 10.1016/j.ejim.2016.02.014] [Medline: 26944564]
64. Chen XS, Wang K. Gender disparities in telehealth use among older adults in the United States during the COVID-19 pandemic. IJPS 2024;10(4):114-124. [doi: 10.36922/ijps.1817]
65. Arcury TA, Sandberg JC, Melius KP, et al. Older adult internet use and eHealth literacy. J Appl Gerontol 2020 Feb;39(2):141-150. [doi: 10.1177/0733464818807468] [Medline: 30353776]
66. Delello JA, McWhorter RR. Reducing the digital divide: connecting older adults to iPad technology. J Appl Gerontol 2017 Jan;36(1):3-28. [doi: 10.1177/0733464815589985] [Medline: 26084479]
67. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253. [doi: 10.2196/17253] [Medline: 32442136]
68. Veras M, Stewart J, Deonandan R, et al. Cost analysis of a home-based virtual reality rehabilitation to improve upper limb function in stroke survivors. GJHS 2020;12(2):98. [doi: 10.5539/gjhs.v12n2p98]
69. Kairy D, Lehoux P, Vincent C, Visintin M. A systematic review of clinical outcomes, clinical process, healthcare utilization and costs associated with telerehabilitation. Disabil Rehabil 2009;31(6):427-447. [doi: 10.1080/09638280802062553] [Medline: 18720118]
70. Houser SH, Flite CA, Foster SL. Privacy and security risk factors related to telehealth services - a systematic review. Perspect Health Inf Manag 2023;20(1):1f. [Medline: 37215337]
71. Mannheim I, Schwartz E, Xi W, et al. Inclusion of older adults in the research and design of digital technology. Int J Environ Res Public Health 2019 Oct 2;16(19):19. [doi: 10.3390/ijerph16193718] [Medline: 31581632]

Abbreviations

MERTH: Metaverse Equitable Rehabilitation Therapy

Edited by C Lee; submitted 05.12.24; peer-reviewed by LK Chua, T Adeniji; revised version received 15.05.25; accepted 03.07.25; published 13.08.25.

Please cite as:
Veras M, Auger LP, Sigouin J, Gheidari N, Nelson MLA, Miller WC, Hudon A, Kairy D
Ethics and Equity Challenges in Telerehabilitation for Older Adults: Rapid Review
JMIR Aging 2025;8:e69660
URL: https://aging.jmir.org/2025/1/e69660
doi:10.2196/69660

Patient-Related Barriers to Digital Technology Adoption in Alzheimer Disease: Systematic Review

Andrea Panzavolta[1], MS; Andrea Arighi[2], MD; Emanuele Guido[1], MD; Luigi Lavorgna[3], MD, PhD; Francesco Di Lorenzo[4], MD, PhD; Alessandra Dodich[5], PhD; Chiara Cerami[1,6], MD

[1]IUSS Cognitive Neuroscience ICoN Center, Scuola Universitaria Superiore IUSS, Piazza della Vittoria, 15, Pavia, Italy

[2]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[3]Division of Neurology, Department of Advanced Medical and Surgical Sciences, AOU-University of Campania, Naples, Italy

[4]Experimental Neuropsychophisiology Unit, Fondazione Santa Lucia IRCCS, Rome, Italy

[5]Center for Mind/Brain Sciences-CIMeC, University of Trento, Rovereto, Italy

[6]Brain e-Health Aging (BeA) Laboratory, Department of Neurorehabilitation, Istituti Clinici Scientifici Maugeri IRCCS, Milan, Italy

Corresponding Author:
Andrea Panzavolta, MS
IUSS Cognitive Neuroscience ICoN Center, Scuola Universitaria Superiore IUSS, Piazza della Vittoria, 15, Pavia, Italy

Abstract

Background: Digital technology in dementia is an area of great development with varying experiences across countries. However, novel digital solutions often lack a patient-oriented perspective, and several relevant barriers prevent their use in clinics.

Objective: In this study, we reviewed the existing literature on knowledge, familiarity, and competence in using digital technology and on attitude and experiences with digital tools in Alzheimer disease. The main research question is whether digital competence and attitudes of patients and caregivers may affect the adoption of digital technology.

Methods: Following the PRISMA guidelines, a literature search was conducted by two researchers in the group. Inter-rater reliability was calculated with Cohen κ statistics. The risk of bias assessment was also recorded.

Results: Of 597 initial records, only 18 papers were considered eligible. Analyses of inter-rater reliability showed good agreement levels. Significant heterogeneity in study design, sample features, and measurement tools emerged across studies. Quality assessment showed a middle-high overall quality of evidence. The main factors affecting the adoption of digital technology in patients and caregivers are severity of cognitive deficits, timing of adoption, and the availability of training and support. Additional factors are age, type of digital device, and ease of use of the digital solution.

Conclusions: Adoption of digital technology in dementia is hampered by many patient-related barriers. Improving digital competence in patient-caregiver dyads and implementing systematic, patient-oriented strategies for the development and use of digital tools are needed for a successful incorporation of digital technology in memory clinics.

(JMIR Aging 2025;8:e64324) doi:10.2196/64324

KEYWORDS

digital technology; digital e-health; accessibility; user-friendliness; neurocognitive disorders; Alzheimer disease; dementia

Introduction

The use of digital technology for the prevention, diagnosis, management, and assistance of patients with cognitive disorders and their caregivers has witnessed a dramatic increase in recent years. Telemedicine and eHealth technology services have proven to be valuable instruments for remote support and care in neurocognitive disorders, as consistently shown during the COVID-19 pandemic (see [1,2] for examples). The severe constraints on health care resources and the reduced time available to chronic patients during the pandemic have demonstrated the urgent need for the reorganization of memory clinic services to ensure adequate diagnosis and care [3]. Accordingly, many initiatives have tested digital technology solutions and telemedicine services in dementia settings [4,5], particularly in Alzheimer disease (AD), which represents one of the highest-risk and fastest-growing burdens on the health care system [6,7]. Literature evidence showed good levels of feasibility and effectiveness, particularly in individuals living in remote or underserved geographical areas or where in-person access to care facilities is limited [8]. Overall, patients with cognitive disorders and caregivers reported good satisfaction rates for digital technology, suggesting a general propensity toward adopting it routinely [8-10].

The literature highlights various solutions, including digital diagnostic instruments, tools for active and passive monitoring, and digital technologies supporting cognitive and motor rehabilitation. These tools have shown promising results not

only in optimizing diagnostic pathways, disease management, and treatments, in avoiding dysfunctional and harmful disease trajectories, but also in helping the empowerment of patients and caregivers through active engagement in their care pathways [11,12]. However, despite the growing evidence suggesting the potential benefit of digital technology tools in cognitive patients, their integration into the daily routines of memory clinics is still limited [13].

Preliminary evidence suggested the presence of enabling factors such as familiarity, acceptability, and a positive attitude toward digital technology that could facilitate its adoption [9,13,14]. Conversely, inadequate experience with technology, poor digital literacy, and low education levels, as well as insufficient accommodation for motor and sensory impairments, may cause reduced acceptability and engagement in older adults [9,15]. However, a clear understanding of factors promoting or hampering the use of digital tools in memory clinics is still lacking.

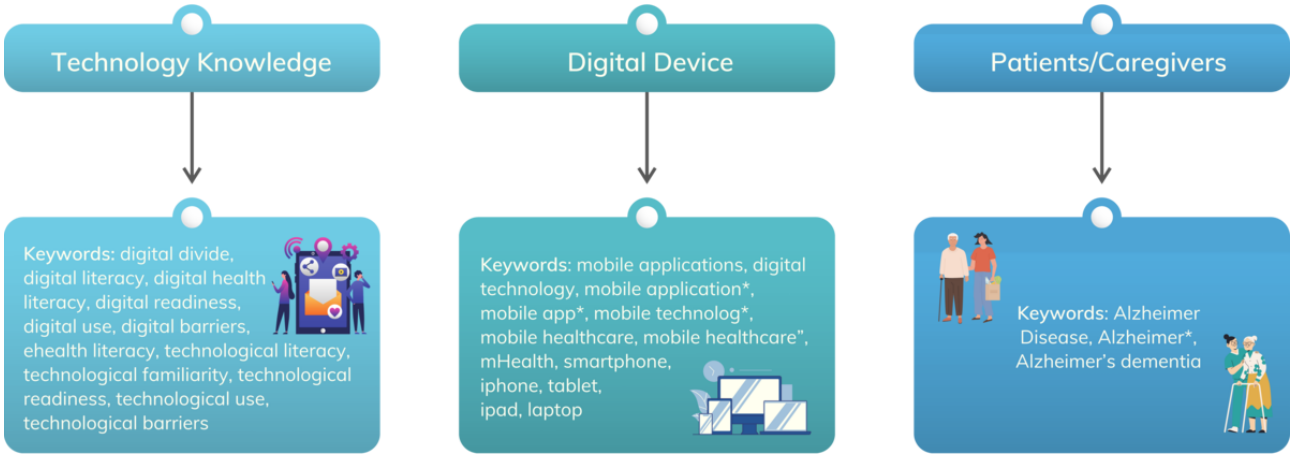
Given these considerations, we reviewed existing literature with a particular interest in studies reporting features facilitating the use of digital technology in users (eg, competence, attitudes) or evaluating the performance of digital solutions (eg, usability). The main research question is whether digital competence and the attitudes of patients and caregivers may affect the adoption of digital technology. The final goal is to explore real-world needs, facilitators, and barriers in the field and to provide patient-oriented guidance for using and implementing novel digital technology tools in memory clinics.

Methods

Search Strategy and Selection Criteria

A systematic search was conducted by two researchers of the group using PubMed, Scopus, Embase, and CINAHL (Cumulated Index in Nursing and Allied Health Literature) databases, following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [16]. According to our research hypothesis, the search strategy included a combination of MeSH (Medical Subject Headings) and relevant text terms focused on assessing familiarity and competence with digital technology and attitudes and experience regarding the use of digital solutions. See Figure 1 and Multimedia Appendix 1 for the full list of search terms. Given our theoretical framework and the growing body of literature on digital technologies in dementia, the search was focused on AD studies, aiming to improve consistency among findings.

Figure 1. Keywords for the literature search.



Inclusion criteria for paper selection were as follows: (1) paper: full original research articles (excluding conference abstracts, case reports, reviews, and book chapters); (2) population of interest: individuals with AD (ranging from mild cognitive impairment to moderate-severe dementia stages) and their caregivers; (3) measurement: papers reporting quantitative and qualitative measures (ie, questionnaires, scales, or interviews) for assessing knowledge or familiarity with digital technology and attitude or experience in using digital solutions; (4) language and time span: papers published in English from 2010 up to October 15, 2023. Articles published before 2010 were excluded due to their reliance on outdated technologies, such as personal digital assistants, which may be less relevant for a contemporary audience.

Study Selection and Data Extraction

The final set of records was uploaded onto Rayyan, a free web and mobile app designed to facilitate the initial screening of abstracts and titles through semiautomation, ensuring a high level of usability [17]. Data extraction was performed encompassing study identifiers (eg, authors, year of publication), sample features (eg, participants, diagnosis, group sizes), description of study methods (eg, type of digital solution and measurement), and findings. To ensure consistency in study selection and data extraction, the Cohen κ statistic was used to assess inter-rater reliability during the screening and eligibility phases [18]. Discrepancies were resolved through discussion, with input from an external senior expert when necessary.

Risk of Bias Assessment

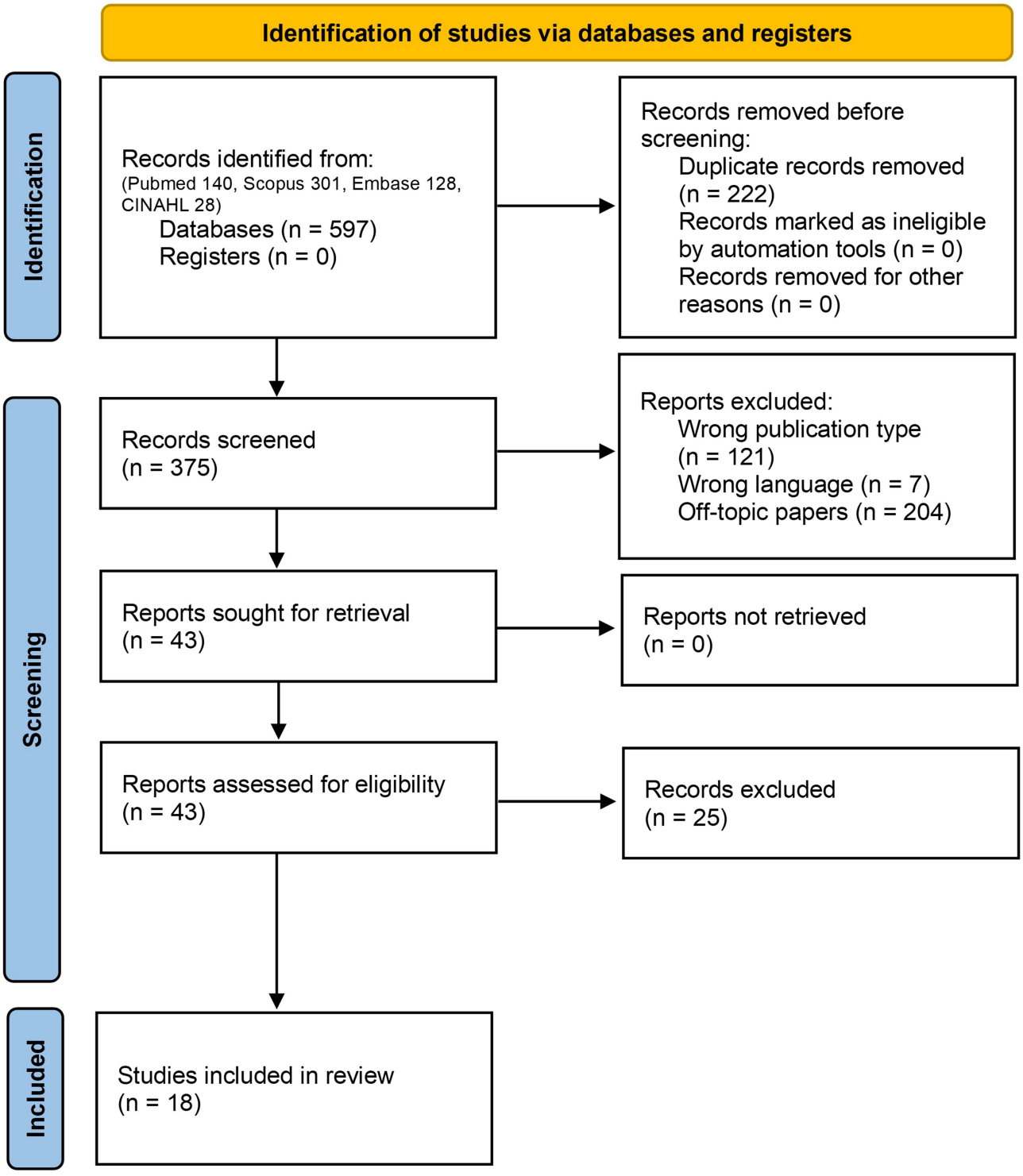
A risk of bias assessment was conducted using the Appraisal tool for Cross-Sectional Studies questionnaire [19] and the National Institutes of Health (NIH) Study Quality Assessment (SQA) Tool [20] for observational cohort and cross-sectional studies. These questionnaires evaluate the quality of human research studies, providing a global score as the rate of positive answers out of either 20 or 14 questions, and indicate overall quality and bias for each study. Considering the percentage of

positive answers, categorical ratings were devised as follows: high quality (>70%), middle (70% - 50%), and low (≤49%) quality.

Results

Out of the 597 articles initially retrieved, 554 papers were excluded during the screening phase (ie, 222 duplicates, 121 nonoriginal papers, 204 off-topic articles, and 7 articles not in English). A detailed review of the full texts led to the exclusion of 25 articles, with 18 papers meeting the eligibility criteria for this study. See Figure 2 for details of the PRISMA flow diagram. Analyses of inter-rater reliability showed good agreement levels for both the screening (κ=0.76; 94.46% agreement) and the eligibility (κ=0.75; 87.76% agreement) phases. Quality assessment of the included studies showed a middle-high quality of evidence.

Figure 2. PRISMA flowchart reporting literature search strategy. CINAHL: Cumulated Index in Nursing and Allied Health Literature.



Retained papers showed high heterogeneity in terms of study design (ie, qualitative vs quantitative), sample features (eg, patients and caregivers with various disease severity), and data type (eg, unstructured vs structured data). Given the high

methodological and population-based heterogeneity across studies, results are presented as a narrative synthesis, focusing on shared themes and distinct sources of variability rather than on statistical aggregation.

See Tables 1 and 2 for details of included papers and the quality assessment.

Table . Summary of papers exploring knowledge, familiarity, and competence with digital technology in patients with Alzheimer disease and caregivers.

Authors and country	Year	Sample	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Guzman-Parra et al [21]; Spain, Sweden	2020	1086 MCI[a] and dementia patient-caregiver dyads	Ad hoc-developed questionnaire to assess use and familiarity with touchscreen mHealth[b] technology (TechPH questionnaire [22])	Low technophilia in patients, young, male, highly educated, living with children, having better mental health	High quality (NIH[c] Study QA[d] Tool Score=75%)
Jacobs et al [23]; US	2021	68 MCI patients 121 ADD[e] patients	Ad hoc-developed web survey	Higher willingness to remote cognitive testing in MCI; significantly less access to video chat–capable technology in ADD; old age and low educational level influence access to technology and willingness to participate in remote cognitive assessment	High quality (NIH Study QA Tool Score=70%)
Christiansen et al [24]; Belgium, Spain, Sweden	2021	1082 MCI patients	Ad hoc-developed questionnaire to assess user experience with mHealth technology	Good to excellent QoL[f] in participants with moderately or high technical skills; variation in technical skills and internet use was a relevant obstacle	High quality (NIH Study QA Tool Score=70%)
Wójcik et al [25]; Poland	2021	102 ADD caregivers	Ad hoc-developed questionnaires to assess acceptance of mHealth technology and laptop/PC use	Age, gender, and education level impact on technology acceptance; old age impact on computer but not on smartphone use; digital technology was perceived as useful for daily caring (locomotion, toileting, and meals)	High quality (NIH Study QA Tool Score=80%)
Albers et al [26]; US	2022	20 people with mild-to-moderate memory complaints patient-caregiver dyads	Semistructured interviews on digital technology use pre- and post-COVID-19 pandemic, adoption of digital technology during the pandemic, facilitators, and barriers to digital technology adoption	Technology use may benefit on maintaining social connections, alleviating boredom, and fostering caregiver relief; patient dependence and low technological literacy prevented the use of digital technology	High quality (NIH Study QA Tool Score=70%)
Talbot and Briggs [27]; UK	2022	19 ADD, mixed and related dementia patients	Semistructured interviews on digital technology use and experience during the COVID-19 pandemic	Low usability and cognitive fatigue in the use of digital technologies; need for training on the use and for a better engagement	Middle quality (NIH Study QA Tool Score=60%)
Nguyen et al [28]; US	2022	124 ADD and related dementia caregivers	Ad hoc-developed web survey to assess experience with iN2L Samsung Galaxy Tablet	Benefit on caregiver wellbeing, alleviating stress, increasing satisfaction, and improving access to supportive programs	Middle quality (NIH Study QA Tool Score=64%)

Authors and country	Year	Sample	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Wilson et al [29]; UK	2023	29 MCI, ADD and related dementia patient-caregiver dyads	Videoconferencing or phone interviews of mhealth technology use	Smart devices are valuable, versatile tools for essential and meaningful activities, and necessary devices to participate in modern life; need for training on the use of smart devices	Middle quality (NIH Study QA Tool Score=50%)

[a]MCI: Mild cognitive impairment.

[b]mHealth: mobile health.

[c]NIH: National Institutes of Health.

[d]QA: Quality assessment.

[e]ADD: Alzheimer disease dementia.

[f]QoL: Quality of Life.

Table . Summary of papers exploring use, acceptability, and usability of digital solutions in AD patients and caregivers.

Authors and country	Year	Sample	Digital solution	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Lim et al [30]; Australia	2013	21 unspecified dementia patient-caregiver dyads	iPad and 11 iPad apps	Ad hoc-developed questionnaires to assess experience, ability to use, engagement and utility	Half of dementia patients were able to engage with and use the iPad independently	Middle quality (NIH[a] Study QA[b] Tool Score=60%)
Zmily et al [31]; Jordan	2014	10 early ADD[c] patients	Samsung Galaxy Tablet and Android app (ADcope)	Ad hoc-developed questionnaires, NASA-TLX[d] index workload assessment [32]	Low workload scores; good post-task satisfaction; successful use even without any prior experience	Middle quality (NIH Study QA Tool Score=50%)
Brown et al [33]; US	2016	11 ADD caregivers	Android app (CareHeroes App)	Ad hoc-developed web-based survey	Perceived easiness to perform tasks despite medium-low proficiency with technology	High quality (NIH Study QA Tool Score=80%)
Killin et al [34]; UK	2018	10 ADD, vascular or mixed dementia patient-caregiver dyads	Digital Support Platform (DSP)	Semistructured interviews	Caregiver use was better than that of patients; High interest in learning to use technology more effectively and enjoyed having their own tablet devices; Need of training in the use of new technology	Middle quality (NIH Study QA Tool Score=50%)
Ruggiano et al [35]; US	2019	36 ADD and related dementia caregivers	Android app (Care IT)	Interviews and focus group	eHealth and individual technologies may not fully meet the needs of caregivers; need for more effective, easy-to-use, and more widely disseminated – especially for caregivers from a disadvantaged background	High quality (NIH Study QA Tool Score=70%)
Øksnebjerg et al [36]; Denmark	2020	112 ADD, vascular dementia patients 98 caregivers	ReACT (Rehabilitation in Alzheimer disease using Cognitive support Technology) app	Ad hoc-developed web-based survey (including USEdem questionnaire)	Need for timely introduction of digital technology; need of caregiver support for the adoption of digital solution	High quality (NIH Study QA Tool Score=85%)
Evans et al [37]; UK	2021	26 ADD and related dementia patient-caregiver dyads	App-based prompter for a touchscreen Tablet	Semistructured interviews on experience and usefulness	Attitudes to technology, perceived utility, and emotional impact of needing help impact the acceptance	High quality (NIH Study QA Tool Score=80%)

Authors and country	Year	Sample	Digital solution	Measurements (qualitative/quantitative)	Study findings	Quality assessment
Berge et al [38]; Norway	2022	24 ADD patient-caregiver dyads	iOS app (Alight)	SUS[e] scale [39] modified; ad hoc-developed questionnaire to explore adoption, user-friendliness, usefulness, and impact	High adoption and feasibility; Need of timely introduction; 50% of the accepting dyads had difficulties independently managing the digital solution	Middle quality (NIH Study QA Tool Score=65%)
Skirrow et al [40]; UK, US	2022	73 MCI/mild[f] ADD patients 78 healthy control subjects	Story recall task app	Ad hoc-developed questionnaires to assess usability and task engagement	Technical problems, easy use of the app, and a broad interest in the tasks; Modest improvement of recall	High quality (NIH Study QA Tool Score=79%)
Rossetto et al [41]; Italy	2023	11 MCI, 19 mild AD patient-caregiver dyads	Telerehabilitation app and web-based software (ABILITY)	SUS scale [39], percentage of sessions attended within 6 weeks (adherence)	High adherence rate, usability, and treatment efficacy on global cognitive level in the digital-treated group	High quality (NIH Study QA Tool Score=93%)

[a]NIH: National Institutes of Health.

[b]QA: Quality assessment.

[c]ADD: Alzheimer disease dementia.

[d]NASA TLX: NASA Task load index.

[e]SUS: System usability scale.

[f]MCI: Mild cognitive impairment.

Evidence From Papers Reporting Knowledge, Familiarity, and Competence With Digital Technology

Among studies investigating factors affecting knowledge, familiarity, and competence in digital mobile health (mHealth) technology use [21,23-29], the degree of cognitive impairment was the first factor significantly affecting digital competence in the adoption of technology solutions. Individuals with milder cognitive deficits showed higher competence in the use of digital devices and the internet and consequently, greater willingness levels to participate in studies compared to those with more advanced cognitive decline [23,24]. Overall, users exhibited greater openness to digital technology when they had prior experience and familiarity and good digital literacy [21,23,24]. Significant challenges were reported in studies related to limited technical skills of both patients and caregivers [26-28]. Nonetheless, while overall caregivers showed greater competence and enthusiasm for digital technology than patients, other factors, as male sex, higher education, and younger age, impact at the individual level on familiarity and willingness to participate to studies [21,25].

Whereas smartphone acceptance is consistent across age groups, computer use declined with age and in patients versus caregivers, indicating an age-related barrier to some technologies [27-29]. For example, the use of tablets represented a significant barrier in the VITAL at Home project [28], in which caregivers required substantial guidance to effectively use the software implemented in tablets.

Evidence From Papers Reporting Acceptability and Usability of Digital Solutions

Studies examining the use of digital solutions among AD patients and caregivers [30,31,33-38,40,41] consistently found that IT system accessibility and ease of use of the proposed solution are crucial factors influencing adoption rates. Ease of use of the digital health solutions and attitude to technology represent key factors enhancing acceptance of digital tools in daily routines [30,31,33,35,37,40]. The likelihood of sustained use is related to the timing of their adoption within the disease course, with earlier implementation correlating with better long-term engagement for both patients and caregivers [34,36]. The provision of informal support and training is another enabling factor ensuring long-term and better use [31,36,41]. As the disease progresses, the severity of cognitive decline and the perceived cognitive workload pose relevant challenges for constant engagement and use [31,37,38].

Discussion

Digital technology in dementia is an area of rapid development marked by diverse approaches across countries. A huge step forward has been made during the COVID-19 pandemic, acting as a catalyst for extensive research and development of eHealth

and mHealth solutions, favoring the testing of digital solutions in memory clinics and promoting their implementation in practice [2,4,42]. In this study, we reviewed the existing literature on digital competence, attitude to technology, and use of digital tools in AD, with the final aim to underline main factors affecting the adoption of digital solutions and provide useful information for their future development and use in memory clinics.

Literature evidence revealed several barriers in adopting digital technology in dementia. Among the main factors, the severity of cognitive deficits is the first key element. As adherence to digital technology interventions decreases with the severity of cognitive decline [23,24,31,37,38], clinicians and researchers should embrace technology-based solutions at an earlier stage of the care pathway with the aim to improve use and effectiveness. With the progression of the disease, poor verbal and motor initiative may affect patient engagement with digital technology, reducing interest, initiative, and participation [27]. In such cases, behavioral or sensory-based interventions, such as sensory stimulation delivered through digital devices, represent an appropriate choice. Therefore, an in-depth assessment of the behavioral and cognitive profile should always guide the specialist to choose the right digital tool to use and address patients accordingly.

Age, digital literacy, and special needs (eg, visual or auditory impairments) are additional variables that limit the access or use of digital technology both for patients and caregivers, without affecting the will to participate in telecare services [21,25,43]. These aspects are of particular relevance in memory clinics, where final users are often patients and caregivers in old adulthood [44]. The type of digital device may also impact the feasibility and real-life implementation of digital protocols, with smartphones being more accessible compared to computers or other devices [27,29,30]. Consequently, smartphones show higher acceptability scores in patients [45,46], although some challenges in the use of touchscreen and software updates are reported, particularly among those with poor digital literacy and visual impairments [24].

Another crucial barrier to the stable use of eHealth and mHealth solutions is the need for training and support [27,29,30,34,36]. The availability of formal support in the technology setup or training has been shown to positively impact willingness and acceptability [30,34,41]. As individuals may be unlikely to embrace burdensome technologies [26], preliminary training at health care facilities before the in-home adoption should be recommended to overcome potential technophobia in patients and caregivers [13,21]. The evaluation of specific digital needs in each single patient-caregiver dyad may guarantee a more inclusive and practical approach.

Therefore, the lack of active involvement of final users in the design, development, and testing of novel digital solutions crucially contributes to the poor applicability of these instruments in real-life scenarios [47].

In addition, developing technologies without user burden is pivotal for their success. Minimizing the burden on caregivers is of paramount importance, as they are often responsible for managing digital tools (eg, calendars for medication or televisits). The mitigation of stress and anxiety due to the use of technologies is crucial for caregivers [26,35,37,38], particularly for those individuals with disadvantageous socioeconomical backgrounds, who face additional challenges in adapting to the tools and using them effectively [35]. Co-design with end-users and adaptation of the digital tool to real-life needs are therefore critical [47].

Several limitations characterize the body of research analyzed in this review. Many studies involved small sample sizes that were not fully representative of the AD population, often excluding individuals who were unable or unwilling to participate for various reasons. Availability of high-speed internet connection [23,24], recruitment or screening modality (eg, through email, social media or videoconference) [26,27], and socioeconomic status of participants [23,37] are additional variables affecting results, while study design (eg, lack of control group or randomized approach [28]) and measurement issues (eg, lack of pre-post measures [36]) are crucial aspects that hamper the reliability of data findings.

In conclusion, digital tools have the potential to significantly influence the quality of dementia services acting on different dimensions [13]. They can increase effectiveness by enabling faster access to specialist care, better diagnoses and treatments, and preventing avoidable hospitalizations. They improve timeliness, reducing waiting lists and unnecessary travel. They ensure patient-centeredness and safe care, providing care tailored to individual needs and values and treating patients more appropriately. They provide integrated, more efficient, and equitable care by taking an interdisciplinary approach involving multiple specialists, improving cost-effectiveness, and overcoming geographic barriers addressing cultural diversity. However, the adoption of digital technology is limited by many patient-related barriers. Improving digital competence in the patient-caregiver dyads and implementing a systematic patient-oriented strategy for the development and use of digital tools (eg, by promoting participated design, early timing of solution adoption and availability of training and technical support) remain critical factors to consider for the successful incorporation of digital eHealth and mHealth solutions and services into future memory clinics.

Data Availability

Data are supplied in supporting files available for download along with the published manuscript.

Authors' Contributions

Conceptualization: CC, AP

Methodology, CC, LL, AP
Software: AP, EG
Supervision: CC
Writing—original draft preparation, CC, EG, AP
Writing—review and editing: CC, AA, LL, FDL, AD, AP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Combination of Medial Subject Headings (MeSH) and text words used for literature search in different databases.
[DOCX File, 22 KB - aging_v8i1e64324_app1.docx]

Checklist 1
PRISMA 2020 checklist for reporting systematic reviews and meta-analyses.
[DOCX File, 24 KB - aging_v8i1e64324_app2.docx]

References

1. Cuffaro L, Di Lorenzo F, Bonavita S, Tedeschi G, Leocani L, Lavorgna L. Dementia care and COVID-19 pandemic: a necessary digital revolution. Neurol Sci 2020 Aug;41(8):1977-1979. [doi: 10.1007/s10072-020-04512-4] [Medline: 32556746]
2. Lukas H, Xu C, Yu Y, Gao W. Emerging telemedicine tools for remote COVID-19 diagnosis, monitoring, and management. ACS Nano 2020 Dec 22;14(12):16180-16193. [doi: 10.1021/acsnano.0c08494] [Medline: 33314910]
3. Cagnin A, Di Lorenzo R, Marra C, et al. Behavioral and psychological effects of coronavirus disease-19 quarantine in patients with dementia. Front Psychiatry 2020;11(578015):578015. [doi: 10.3389/fpsyt.2020.578015] [Medline: 33033486]
4. Owens AP, Ballard C, Beigi M, et al. Implementing remote memory clinics to enhance clinical care during and after COVID-19. Front Psychiatry 2020;11:579934. [doi: 10.3389/fpsyt.2020.579934] [Medline: 33061927]
5. Tu CC, Weng SY, Hsieh NC, Cheng WC, Alizargar J, Chang KS. Increasing use of telemedicine for neurological disorders during the COVID-19 pandemic: a mini-review. J Multidiscip Healthc 2023;16:411-418. [doi: 10.2147/JMDH.S390013] [Medline: 36820220]
6. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2(1):9. [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
7. Staffaroni AM, Tsoy E, Taylor J, Boxer AL, Possin KL. Digital cognitive assessments for dementia. Pract Neurol (Fort Wash Pa) 2020(24) [FREE Full text] [Medline: 33927583]
8. Sekhon H, Sekhon K, Launay C, et al. Telemedicine and the rural dementia population: a systematic review. Maturitas 2021 Jan;143:105-114. [doi: 10.1016/j.maturitas.2020.09.001] [Medline: 33308615]
9. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
10. Mendez KJW, Labrique AB, Budhathoki C, et al. Use of mobile applications and health technologies among dementia caregivers with chronic conditions: a cross-sectional study. Digit Health 2023;9(20552076231181213):20552076231181213. [doi: 10.1177/20552076231181213] [Medline: 37361436]
11. Yousaf K, Mehmood Z, Awan IA, et al. A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD). Health Care Manag Sci 2020 Jun;23(2):287-309. [doi: 10.1007/s10729-019-09486-0] [Medline: 31218511]
12. Bradley L, Shanker S, Murphy J, Fenge LA, Heward M. Effectiveness of digital technologies to engage and support the wellbeing of people with dementia and family carers at home and in care homes: a scoping review. Dementia (London) 2023 Aug;22(6):1292-1313. [doi: 10.1177/14713012231178445] [Medline: 37235791]
13. Angelopoulou E, Papachristou N, Bougea A, et al. How telemedicine can improve the quality of care for patients with Alzheimer's disease and related dementias? A narrative review. Medicina (Kaunas) 2022 Nov 22;58(12):1705. [doi: 10.3390/medicina58121705] [Medline: 36556907]
14. Stronge AJ, Rogers WA, Fisk AD. Human factors considerations in implementing telemedicine systems to accommodate older adults. J Telemed Telecare 2007;13(1):1-3. [doi: 10.1258/135763307779701158] [Medline: 17288650]
15. Bossen AL, Kim H, Williams KN, Steinhoff AE, Strieker M. Emerging roles for telemedicine and smart technologies in dementia care. Smart Homecare Technol Telehealth 2015;3(49–57):49-57. [doi: 10.2147/SHTT.S59500] [Medline: 26636049]
16. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097. [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
18. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;22(3):276-282. [doi: 10.11613/BM.2012.031] [Medline: 23092060]
19. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 2016 Dec 8;6(12):e011458. [doi: 10.1136/bmjopen-2016-011458] [Medline: 27932337]
20. Study quality assessment tools. NHLBI, NIH. 2021. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [accessed 2023-04-12]
21. Guzman-Parra J, Barnestein-Fonseca P, Guerrero-Pertiñez G, et al. Attitudes and use of information and communication technologies in older adults with mild cognitive impairment or early stages of dementia and their caregivers: cross-sectional study. J Med Internet Res 2020 Jun 1;22(6):e17253. [doi: 10.2196/17253] [Medline: 32442136]
22. Anderberg P, Eivazzadeh S, Berglund JS. A novel instrument for measuring older people's attitudes toward technology (TechPH): development and validation. J Med Internet Res 2019 May 23;21(5):e13951. [doi: 10.2196/13951] [Medline: 31124467]
23. Jacobs DM, Peavy GM, Banks SJ, Gigliotti C, Little EA, Salmon DP. A survey of smartphone and interactive video technology use by participants in Alzheimer's disease research: implications for remote cognitive assessment. Alzheimers Dement (Amst) 2021;13(1):e12188. [doi: 10.1002/dad2.12188] [Medline: 34027018]
24. Christiansen L, Sanmartin Berglund J, Anderberg P, et al. Associations between mobile health technology use and self-rated quality of life: a cross-sectional study on older adults with cognitive impairment. Gerontol Geriatr Med 2021;7:23337214211018924. [doi: 10.1177/23337214211018924] [Medline: 34104685]
25. Wójcik D, Szczechowiak K, Konopka P, et al. Informal dementia caregivers: current technology use and acceptance of technology in care. Int J Environ Res Public Health 2021 Mar 19;18(6):3167. [doi: 10.3390/ijerph18063167] [Medline: 33808644]
26. Albers EA, Mikal J, Millenbah A, et al. The use of technology among persons with memory concerns and their caregivers in the United States during the COVID-19 pandemic: qualitative study. JMIR Aging 2022 Mar 17;5(1):e31552. [doi: 10.2196/31552] [Medline: 35134748]
27. Talbot CV, Briggs P. The use of digital technologies by people with mild-to-moderate dementia during the COVID-19 pandemic: a positive technology perspective. Dementia (London) 2022 May;21(4):1363-1380. [doi: 10.1177/14713012221079477] [Medline: 35333111]
28. Nguyen LT, Prophater LE, Fazio S, et al. Project VITAL at home: impact of technology on loneliness and well-being of family caregivers of people with dementia. Clin Gerontol 2024;47(1):149-160. [doi: 10.1080/07317115.2022.2123726] [Medline: 36111824]
29. Wilson SA, Byrne P, Rodgers SE. "I'd be lost without my smartphone": a qualitative analysis of the use of smartphones and tablets by people living with dementia, mild cognitive impairment, and their caregivers. Aging Ment Health 2024 Apr;28(4):595-603. [doi: 10.1080/13607863.2023.2205585] [Medline: 37139944]
30. Lim FS, Wallace T, Luszcz MA, Reynolds KJ. Usability of tablet computers by people with early-stage dementia. Gerontology 2013;59(2):174-182. [doi: 10.1159/000343986] [Medline: 23257664]
31. Zmily A, Mowafi Y, Mashal E. Study of the usability of spaced retrieval exercise using mobile devices for Alzheimer's disease rehabilitation. JMIR Mhealth Uhealth 2014 Aug 14;2(3):e31. [doi: 10.2196/mhealth.3136] [Medline: 25124077]
32. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. Human mental workload 1988:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
33. Brown EL, Ruggiano N, Page TF. CareHeroes web and android. Res Gerontol Nurs 2016;9(4):193-203. [doi: 10.3928/19404921-20160229-02] [Medline: 29977440]
34. Killin LOJ, Russ TC, Surdhar SK, et al. Digital Support Platform: a qualitative research study investigating the feasibility of an internet-based, postdiagnostic support platform for families living with dementia. BMJ Open 2018 Apr 12;8(4):e020281. [doi: 10.1136/bmjopen-2017-020281] [Medline: 29654028]
35. Ruggiano N, Brown EL, Shaw S, et al. The potential of information technology to navigate caregiving systems: perspectives from dementia caregivers. J Gerontol Soc Work 2019;62(4):432-450. [doi: 10.1080/01634372.2018.1546786] [Medline: 30422754]
36. Øksnebjerg L, Woods B, Ruth K, et al. A tablet app supporting self-management for people with dementia: explorative study of adoption and use patterns. JMIR Mhealth Uhealth 2020 Jan 17;8(1):e14694. [doi: 10.2196/14694] [Medline: 31951217]
37. Evans N, Boyd H, Harris N, et al. The experience of using prompting technology from the perspective of people with Dementia and their primary carers. Aging Ment Health 2021 Aug;25(8):1433-1441. [doi: 10.1080/13607863.2020.1745145] [Medline: 32223428]
38. Berge LI, Gedde MH, Torrado Vidal JC, et al. The acceptability, adoption, and feasibility of a music application developed using participatory design for home-dwelling persons with dementia and their caregivers. The "Alight" app in the LIVE@Home.Path trial. Front Psychiatry 2022;13:949393. [doi: 10.3389/fpsyt.2022.949393] [Medline: 36061298]

39. Brooke J. SUS: A quick and dirty usability scale. Usability Eval Ind 1995 Nov 30 [FREE Full text]
40. Skirrow C, Meszaros M, Meepegama U, et al. Validation of a remote and fully automated story recall task to assess for early cognitive impairment in older adults: longitudinal case-control observational study. JMIR Aging 2022 Sep 30;5(3):e37090. [doi: 10.2196/37090] [Medline: 36178715]
41. Rossetto F, Isernia S, Realdon O, et al. A digital health home intervention for people within the Alzheimer's disease continuum: results from the Ability-TelerehABILITation pilot randomized controlled trial. Ann Med 2023 Dec;55(1):1080-1091. [doi: 10.1080/07853890.2023.2185672] [Medline: 36929703]
42. León-Salas B, González-Hernández Y, Infante-Ventura D, et al. Telemedicine for neurological diseases: a systematic review and meta-analysis. Eur J Neurol 2023 Jan;30(1):241-254. [doi: 10.1111/ene.15599] [Medline: 36256522]
43. Conway A, Ryan A, Harkin D, Mc Cauley C, Goode D. A review of the factors influencing adoption of digital health applications for people living with dementia. Digit Health 2023;9:20552076231162985. [doi: 10.1177/20552076231162985] [Medline: 36937696]
44. Canevelli M, Di Pucchio A, Marzolini F, et al. A national survey of centers for cognitive disorders and dementias in Italy. J Alzheimers Dis 2021;83(4):1849-1857. [doi: 10.3233/JAD-210634] [Medline: 34459403]
45. Joe J, Demiris G. Older adults and mobile phones for health: a review. J Biomed Inform 2013 Oct;46(5):947-954. [doi: 10.1016/j.jbi.2013.06.008] [Medline: 23810858]
46. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038. [doi: 10.1093/geroni/igy038] [Medline: 30619948]
47. Bitkina O, Kim HK, Park J. Usability and user experience of medical devices: an overview of the current state, analysis methodologies, and future challenges. Int J Ind Ergon 2020 Mar;76:102932. [doi: 10.1016/j.ergon.2020.102932]

Abbreviations

AD: Alzheimer disease
CINAHL: Cumulated Index in Nursing and Allied Health Literature
NIH: National Institutes of Health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SQA: Study Quality Assessment

Edited by AF Abdulai; submitted 15.07.24; peer-reviewed by F Yu, L Cuffaro, RS Gomaa Mahmoud; revised version received 09.12.24; accepted 01.01.25; published 10.04.25.

Please cite as:
Panzavolta A, Arighi A, Guido E, Lavorgna L, Di Lorenzo F, Dodich A, Cerami C
Patient-Related Barriers to Digital Technology Adoption in Alzheimer Disease: Systematic Review
JMIR Aging 2025;8:e64324
URL: https://aging.jmir.org/2025/1/e64324
doi:10.2196/64324

Review

Best Practices for Implementing Electronic Care Records in Adult Social Care: Rapid Scoping Review

Martha Snow[1], MA; Wagner Silva-Ribeiro[1], PhD; Mary Baginsky[2], PhD; Sonya Di Giorgio[3], MA; Nicola Farrelly[4], PhD; Cath Larkins[4], PhD; Karen Poole[3], MSc; Nicole Steils[2], PhD; Joanne Westwood[4], PhD; Juliette Malley[1], PhD

[1]Care Policy and Evaluation Centre, London School of Economics and Political Science (LSE), London, United Kingdom

[2]NIHR Health & Social Care Workforce Research Unit, King's College London, London, United Kingdom

[3]Libraries & Collections, King's College London, London, United Kingdom

[4]School of Health, Social Work and Sport, University of Central Lancashire, Preston, Lancashire, United Kingdom

Corresponding Author:
Martha Snow, MA
Care Policy and Evaluation Centre
London School of Economics and Political Science (LSE)
PAN 8.01J
Houghton Street
London, WC2A 2AE
United Kingdom
Phone: 44 (0)20 7405 7686
Email: m.snow@lse.ac.uk

Abstract

Background: In the past decade, the use of digital or electronic records in social care has risen worldwide, capturing key information for service delivery. The COVID-19 pandemic accelerated digitization in health and social care. For example, the UK government created a fund specifically for adult social care provider organizations to adopt digital social care records. These developments offer valuable learning opportunities for implementing digital care records in adult social care settings.

Objective: This rapid scoping review aimed to understand what is known about the implementation of digital care records in adult social care and how implementation varies across use cases, settings, and broader contexts.

Methods: A scoping review methodology was used, with amendments made to enable a rapid review. Comprehensive searches based on the concepts of digital care records, social care, and interoperability were conducted across the MEDLINE, EmCare, Web of Science Core Collection, HMIC Health Management Information Consortium, Social Policy and Practice, and Social Services Abstracts databases. Studies published between 2018 and 2023 in English were included. One reviewer screened titles and abstracts, while 2 reviewers extracted data. Thematic analysis mapped findings against the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework.

Results: Our search identified 2499 references. After screening titles and abstracts, 71 records were selected for full-text review, resulting in 31 references from 29 studies. Studies originated from 11 countries, including 1 multicountry study, with the United Kingdom being the most represented (10/29, 34%). Studies were most often conducted in nursing homes or facilities (7/29, 24%) with older people as the target population (6/29, 21%). Health records were the most investigated record type (12/29, 41%). We identified 45 facilitators and 102 barriers to digital care record implementation across 28 studies, spanning 6 of the 7 NASSS framework domains and aligning with 5 overarching themes that require greater active management regarding implementation. Intended or actual implementation outcomes were reported in 17 (59%) of the 29 studies.

Conclusions: The findings suggest that implementation is complex due to a lack of consensus on what digital care records and expected outcomes and impacts should look like. The literature often lacks clear definitions and robust study designs. To be successful, implementation should consider complexity, while studies should use robust frameworks and mixed methods or quantitative designs where appropriate. Future research should define the target population, gather data on carer or service user experiences, and focus on digital care records specifically used in social care.

(JMIR Aging 2025;8:e60107) doi:10.2196/60107

KEYWORDS

digital care records; adult social care; digitization; domiciliary care; care homes; electronic care records; PRISMA

Introduction

Background

The demand for adult social care is vast. Global demographic changes throughout the 20th century have led to substantial population aging, decreased mortality and communicable diseases, and increased chronic noncommunicable diseases. Consequently, more adults and older people have long-term care needs, particularly in high-income countries where the epidemiological transition began earlier [1]. Technology has been proposed to help manage this increasing demand in health and social care by improving efficiency, care quality, and effectiveness [2-4]. Digital care records are one such innovation in adult social care.

In this paper, adult social care refers to long-term, aged, or disability care, including care homes; support in the home; domiciliary care (eg, personal care, practical tasks, and crisis support); community-based support such as inclusive arts programs; and social relationships that aim to keep people independent, active, and living well.

The use of digital care records has increased across various adult social care settings and countries since 2012 [5]. These records capture key information for service delivery, including individuals' characteristics, the care they receive, and how they respond to it. They monitor service users and track service delivery, supporting care planning, medication, and assessments [6-10]. In addition, they serve administrative purposes [8,11,12], support compliance with data documentation regulations [13], and inform care delivery decisions [14,15]. Different terms are used to describe digital care records in social care. In the United Kingdom, the term digital social care records (DSCRs) is common. In North America, parts of Europe, and Australia, terms such as electronic health records [16-20], electronic patient records [6], or electronic medical records [9,19] are often used. Digital care records can be part of health information exchange initiatives, which facilitate data sharing across health and social care to improve care continuity and efficiency [16-20].

Despite the increasing use of digital care records, much of the existing literature focuses on their implementation in nursing homes or approaches the topic from a social work perspective, failing to capture the full scope of adult social care. One systematic review of electronic health records identified that health information exchange is facilitated by workflow integration and flexible organizational culture and impeded by incomplete data, inefficiency, and unfavorable market conditions [21]. Another review found that electronic health records support health outcomes, clinical documentation management, and decision-making [2]. To the best of our knowledge, the only previous review of DSCRs that looked at the benefits of implementation was a review by Greenstock [22]. This literature review highlighted improved documentation and health outcomes as well as increased collaboration and communication, efficiency, quality of care, client or family involvement, and risk management [22]. It identified less evidence regarding financial benefits and increased workforce satisfaction [22]. It is unclear how many benefits were realized versus anticipated [22]. A scoping review of electronic information systems in social care found that they can negatively affect social workers' priorities and do not meet sector needs [23].

These reviews predate the COVID-19 pandemic, which accelerated digital system development in health and social care [24]. For example, the UK government injected funds during the COVID-19 pandemic to drive digitization and has since continued these efforts. A specific fund for adult social care provider organizations supports DSCR adoption, with the most recent government target of 80% adoption across adult social care provider organizations in England by March 2025.

Objectives

The intensity of the activity discussed above presented an opportunity to learn about DSCR implementation and impact through evaluation. Considering recent developments, this rapid scoping review sought to assess what is known about DSCR implementation in adult social care settings and identify evidence gaps to inform a rapid evaluation of DSCR implementation. While this purpose has influenced decisions around the methods, such as a rapid approach and more intensive searching for UK literature, the review considers the international literature on DSCR implementation and draws out implications for an international audience.

We mapped our findings against the nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework by Greenhalgh et al [25]. Designed in 2017 as an evidence-based, theory-informed, and pragmatic tool, it helps predict and evaluate the success of technology-supported health or social care programs. As it focuses on adoption, nonadoption, and abandonment of technologies as well as the challenges associated with the scale-up, spread, and maintenance of digital systems, it was deemed appropriate for capturing the field's complexity. The framework was particularly useful during data analysis. Most of the literature retrieved identified large numbers of facilitators of and barriers to DSCR adoption. The NASSS framework helped to position these within an interrelated system and organize them in a way that could provide guidance in areas requiring active management of complexity. As the NASSS framework has been applied more often to health care settings, this review was also an opportunity to explore its value for technology adoption in social care.

Methods

Overview

The rapid scoping review followed the 6-stage framework outlined by Arksey and O'Malley [26], which was later refined by Levac et al [27] and the Joanna Briggs Institute [28]. Following the study by Tricco et al [29], we made some amendments to enable a rapid review. The review is reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for

Scoping Reviews) guidelines (Multimedia Appendix 1) [30]. The search strategy is reported in accordance with the PRISMA-S (Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension) checklist [31]. A protocol for this review was developed using the PRISMA-ScR and registered prospectively with the Open Science Framework on August 9, 2023 [32].

Identifying the Research Question

We used the Joanna Briggs Institute's population, concept, and context framework [28] to formulate the following scoping review questions: (1) What is known about the implementation of DSCRs in social care settings? and (2) How does implementation vary across use cases, social care settings, and the broader context? The subquestions were as follows:

- What DSCR is being used?
- What situation or setting is the DSCR being used in, and which actors are involved?
- What is the broader context within which DSCRs are being implemented or used?
- What is the use case for the DSCR, and what are the intended outcomes and benefits?
- How has the implementation of DSCRs been evaluated or researched, and what theoretical framings have been used?
- What are the intended or actual outcomes and benefits of DSCR implementation?
- What helps or gets in the way of the implementation of DSCRs?

Identifying Relevant Studies

A librarian with experience in undertaking reviews (KP) designed the search in consultation with the research team. The search was undertaken between August 2, 2023, and August 11, 2023, by 2 librarians (KP and SDG) across MEDLINE (through Ovid; KP), EmCare (through Ovid; SDG), Web of Science Core Collection (Clarivate; KP), HMIC Health Management Information Consortium (through Ovid; KP), Social Policy and Practice (through Ovid; KP), and Social Services Abstracts (through ProQuest; SDG) databases.

The search strategy used 3 concepts: digital care records, social care, and interoperability. These concepts were combined in the search string as (Digital Care Records AND Social Care) OR (Social Care AND Interoperability). The interoperability concept was included, as it is central to policy narratives surrounding the implementation of DSCRs in England, with expectations that DSCRs will facilitate data sharing with general practitioners and hospitals. The initial search strategy was developed on MEDLINE (Ovid) by one of the librarians (KP) and run in each database by KP and SDG. Publications were limited to those published in or after 2018 until 2023. The results were limited to the English language. The databases were searched using keywords and controlled vocabulary (eg, Medical Subject Headings or Emtree) where appropriate and adapted according to the requirements of each database. The full search strategy for each database can be found in Multimedia Appendix 2.

There were 3466 results in total. The results were exported to EndNote (Clarivate), and 993 duplicates were removed following a structured process [33], leaving 2473 unique results. These were exported as a research information systems file to Covidence (Veritas Health Innovation Ltd) software [34] for title and abstract screening as well as for full-text review.

In addition, given the intention of informing an evaluation in the context of the English language, we searched key English websites to capture gray literature not identified through the databases. The chosen websites were the Local Government Association [35]; King's Fund [36]; Social Care Institute of Excellence [37]; Centre for Care [38]; Digital Care Hub, formerly Digital Social Care [39]; and TEC Service Association [40]. Searches were also performed on Google, and we contacted experts identified through the review. From these searches and reference checking, 27 references were identified. Of these 27 references, 1 (4%) duplicate was removed, and 1 (4%) reference that reported results from a study already included was merged with the main reference. One reference recommended by an expert was also included. This resulted in 26 references retrieved through our gray literature search.

Study Selection

We included studies that (1) took place within adult social care settings; (2) involved the implementation of a DSCR, which may be referred to by other labels, such as electronic care records and electronic information systems; (3) were carried out using any study design (eg, experimental, quasi-experimental, and observational, including quantitative and qualitative studies); and (4) were published from 2018 onward. This decision was made on the basis that existing reviews have captured the literature on DSCRs up until the end of 2017.

Following rapid review methodology guidance [29,41], all references retrieved from our search were screened by a reviewer with expertise in systematic reviews (WSR). Initial screening was based on titles and abstracts. References were selected for full-text review if they met our inclusion criteria or if it was unclear that they did. The same reviewer (WSR) performed the full-text review. A second reviewer, who is an expert in adult social care research (JM), cross-checked references that were excluded in this phase. Disagreement was discussed until a consensus was reached.

Charting the Data

A data extraction template was developed by the team using Excel (Microsoft Corp). The form included key characteristics of included studies, such as the population, concept, context, study design, and methods, and key findings that were relevant to the review questions. In total, 2 reviewers (MS and WSR) performed the data extraction. Due to the heterogeneity of studies and following best practice, the extraction form was piloted and iteratively adapted through discussions between the 2 reviewers and a third reviewer (JM), who oversaw the extraction process.

During the data extraction, we discovered that 2 publications [42,43] reported results from the same study. Another publication [44] was a preprint version of 1 peer-reviewed article [45], which was also included in the review. All publications were included to ensure we used the information available, but

to avoid duplication of information, we extracted information at the study level rather than the publication level.

Collating, Reporting, and Summarizing the Results

There were several steps involved in collating and reporting the results. We first created a summary of the included studies, categorizing the papers according to relevant study characteristics, such as study design, population, context, methods of data collection and analysis, and theoretical perspectives. We then worked inductively to identify intended or actualized benefits and outcomes and barriers to and facilitators of implementation raised in the papers. Using thematic analysis, we compiled a descriptive overview of the unique barriers and facilitators identified in the papers, including frequency distributions.

We then used the NASSS framework as a sensitizing framework and worked deductively to ensure we had not missed anything of relevance to the NASSS domains. In this process, further barriers and facilitators were identified, and these were mapped alongside those identified from the inductive process to the NASSS domains and subdomains. Where a category was associated with >1 NASSS domain, it was mapped against the domain perceived as most affected.

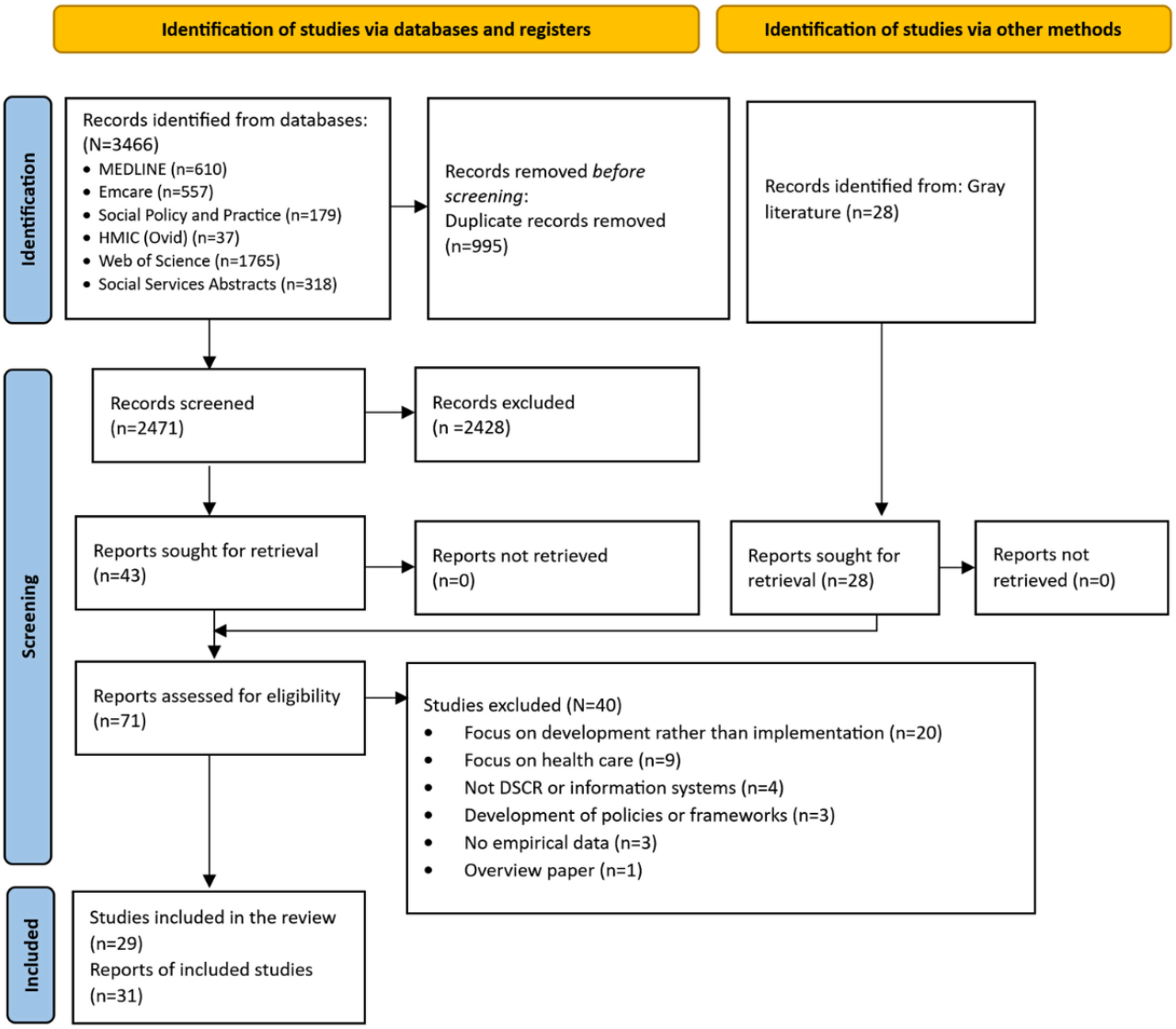
To synthesize our findings, we then grouped the barriers and facilitators into themes capturing complex aspects of the adoption process. Complexity was determined using the NASSS framework, which defines implementation as simple (ie, few components and predictable), complicated (ie, many components but still largely predictable), or complex (ie, many components interacting in a dynamic and unpredictable way) [46]. The more complexity there is in the system, the less likely the technology is to achieve sustained adoption across the system, and the more likely it is to be abandoned [46]. The themes draw attention to areas that require greater active management with respect to implementation [25].

Results

Overview

Our search resulted in 2471 references after duplicates were removed. An additional 28 references were identified through the gray literature search, resulting in 2499 references. After screening references based on titles and abstracts, 71 records were selected for full-text review, of which 31 references were included from 29 different studies (ie, 2 pairs of papers reported on the same studies). The article selection process and reasons for exclusion are presented in Figure 1 [47].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flowchart. DSCR: digital social care record.



Study Characteristics

As Table 1 presents, of the 29 studies, 10 (34%) were conducted in the United Kingdom—7 (24%) in England [3,7,10,20,48-50], 2 (7%) in Scotland [43,45], and 1 (3%) [51] in multiple UK countries. Of the 29 studies, 5 (17%) were conducted in the United States [9,14-16,18], 3 (10%) in Finland [52-54], 2 (7%) in Australia [11,13], 2 (7%) in Canada [19,55], and 2 (7%) in Sweden [12,56]. The remaining studies were conducted in Switzerland [8], Japan [57], Austria [17], and Italy [58] or involved multiple countries [6].

Studies in the United Kingdom were conducted in care homes (3/29, 10%) [43,45,48], health and social care provider organizations (4/29, 14%) [3,7,20,49], multiple social care settings (1/29, 3%) [51], councils (1/29, 3%) [50], and a continuing health care team (1/29, 3%) [10]. More than one-third of studies from other countries were conducted in nursing homes and facilities (7/29, 24%) [6,8,9,14,15,17,18]. Other settings included home care (5/29, 17%) [12,55-58], care homes (1/29, 3%) [13], long-term care facilities (1/29, 3%) [11], an acute care hospital and its neighboring long-term care home (1/29, 3%) [19], assisted living communities (1/29, 3%) [16], health centers in Finland (2/29, 7%) [52,53], and social care services (1/29, 3%) [54].

Table 1. Key characteristics of the included studies (N=29).

Characteristic	Studies, n (%)
Design	
Cohort	1 (3)
Cross-sectional	6 (21)
Mixed methods	8 (28)
Qualitative	14 (48)
Country	
Australia	2 (7)
Austria	1 (3)
Canada	2 (7)
Finland	3 (10)
Italy	1 (3)
Japan	1 (3)
Sweden	2 (7)
Switzerland	1 (3)
United Kingdom	10 (35)
United States	5 (17)
Belgium, Czech Republic, and Spain	1 (3)
Aim	
Barriers and facilitators	9 (31)
Prevalence of use of DSCRs[a]	7 (24)
Professionals' perceptions about DSCRs	4 (14)
Impact of DSCRs on professionals' work	4 (14)
Services' readiness to implement DSCRs	3 (10)
Strategies to improve DSCRs	1 (3)
Impact of DSCRs on health outcomes	1 (3)
Setting	
Assisted living community	1 (3)
Care homes	4 (14)
Continuing or community health care	3 (10)
Councils	1 (3)
Home care	5 (17)
Multisector (ie, health and social care)	6 (21)
Municipal social services	1 (3)
Nursing homes or facilities	7 (24)
Social care provider organizations	1 (3)
Population	
Older people	6 (21)
Adults with care needs	2 (7)
People with dementia	1 (3)
Not specified	20 (69)
Respondents[b]	

Characteristic	Studies, n (%)
Social care staff	10 (35)
Health care staff	8 (28)
Social care managers	8 (28)
Health care managers	3 (10)
Users or carers	2 (7)
Regional stakeholders	2 (7)
National stakeholders	1 (3)
Technology providers or vendors	1 (3)
Not specified	6 (21)
Types of technology	
Health records	12 (41)
Health and social care records	8 (28)
Social care records	4 (14)
Interoperability	4 (14)
Not specified	1 (3)
Theoretical framework used[b]	
The DeLone and McLean model of information systems success	1 (3)
The Wang and Strong quality framework	1 (3)
Activity theory	1 (3)
Sociotechnical systems theory	1 (3)
Computer-supported cooperative work	1 (3)
Design thinking	1 (3)
Nolan stage model	1 (3)
Normalization process theory	2 (7)
Implementation process framework	1 (3)
Unified theory of acceptance and use of technology	1 (3)
Not specified or applicable	20 (69)

[a]DSCR: digital social care record.

[b]Total is >100% because some studies collected information with different types of informants.

Studies investigated several different types of digital care records—from health information technology in general (4/29, 14%) [18,48,55,57] to electronic medical records or electronic health records specifically (12/29, 41%) [6,8,9,11,12,14-18,53,56]. Some were specific systems commissioned by or developed for care providers, local authorities, or regions, such as the aged care ecosystem [13], the Edotto regional information system [58], CareFirst [10], the PASSsystem [51], or CareCentric [3], among others. Of the 29 studies, 4 (14%) focused on interoperability [10,19,20,57]. Studies can be grouped into those which aimed to (1) identify barriers or facilitators to the implementation of DSCRs and information exchange systems (9/29, 31%) [3,10,13,19,20,48-50,53], (2) assess the proportion and prevalence of services using DSCRs and information exchange systems and how these are being used (7/29, 24%) [14-18,45,56], (3) investigate how digital systems affect the work of care professionals or care providers (4/29, 14%) [7,11,51,54], (4) assess care professionals' perceptions about the use of digital systems (4/29, 14%) [6,8,52,55], (5) map services' readiness or maturity and care professionals' capability to adopt DSCRs and information exchange systems (3/29, 10%) [9,12,42], (6) assess potential strategies to improving existing DSCRs or information exchange systems (1/29, 3%) [58], and (7) assess the impact of use of information communication technology (ICT) on health outcomes (1/29, 3%) [57].

To achieve these aims, most studies used a qualitative design (14/29, 48%) [3,6,7,11-13,17,48-50,52,53,56,58]. Of the 29 studies, 8 (28%) used mixed methods [9,10,19,20,45,51,54,55] and 7 (24%) used a quantitative design [8,14-16,18,43,57]. None of the studies that aimed to identify barriers or facilitators, to investigate how digital systems affect work routines, or to assess potential strategies to improve digital systems used quantitative methods. Among the studies that aimed to assess professionals' perceptions of digital systems, only 1 was quantitative [8]. Of

the studies that aimed to assess the proportion and prevalence of digital systems, most (4/29, 14%) used quantitative methods [14-16,18]. The study assessing the impact of ICT on health outcomes was also quantitative [57]. The studies that aimed to map services' readiness and maturity varied between qualitative [12], quantitative [42], and mixed methods approaches [9].

Most studies (20/29, 69%) did not use a theoretical framework to interpret their results [3,8-10,14-17,19,42,45,48-52,54-57]. Of those that did, theories included normalization process theory (2/29, 7%) [20,53], sociotechnical systems theory (1/29, 3%) [6], the Nolan stage model (1/29, 3%) [18], an implementation process framework (1/29, 3%) [13], the DeLone and McLean model of information systems success and the Wang and Strong quality framework (1/29, 3%) [58], activity theory (1/29, 3%) [11], design thinking (1/29, 3%) [12], computer-supported cooperative work (1/29, 3%) [7], and the unified theory of acceptance and use of technology (1/29, 3%) [20].

A detailed list of the characteristics of all studies included in this review is provided in Multimedia Appendix 3 [3,6-20,42-45,48-58].

Summary of Facilitators of and Barriers to the Implementation of DSCRs

Overview

Of the 29 studies, 28 (97%) identified 45 facilitators of and 102 barriers to digital implementation. These were then coded into 32 categories that aligned with the NASSS framework domains: 18 contained facilitators and 24 contained barriers (the total is >32 because some categories contained both facilitators and barriers). The most frequent barriers were related to the digital system lacking interoperability, which was found in 10 of the 29 studies (34%). They also related to insufficient funding or financial incentives and high costs of implementation (9/29, 31%), and technology not matching the context of use (9/29, 31%). Most facilitators were associated with building interorganizational trust and collaborative relationships (5/29, 17%); adequate training (5/29, 17%); anticipating, frontloading, and resourcing the work required to clarify information governance (4/29, 14%); skillful leadership enhancing an organization's digital readiness and capacity for change (4/29, 14%); and high usability of the digital system (4/29, 14%).

Regarding the NASSS framework domains, most facilitators were related to the organization (24/45, 53%). This was followed by the adopter system (8/45, 18%), the technology (6/45, 13%), the value proposition (4/45, 9%), the interaction between domains and adaptation over time (2/45, 4%), and the wider context (1/45, 2%). Most barriers were also related to the organization (52/102, 51%). This was followed by the technology (25/102, 24.5%), the wider context (14/102, 13.7%), the value proposition (6/102, 5.9%), the adopter system (4/102, 3.9%), and the interaction between domains and adaptation over time (1/102, 1%). No barriers or facilitators were related to the condition domain.

The categories containing facilitators and barriers were then organized into five broad themes: (1) the legal and institutional context for holding and sharing data and its effect on the ability and willingness to share data, (2) digital readiness and organizational capacity for change, (3) using and sharing recorded information within technical constraints, (4) alignment between care practices and digital recording practices, and (5) differences between what is expected and what is achievable with digital systems.

A summary of how barriers and facilitators identified in each study were mapped to categories, themes, and the NASSS framework domains and subdomains is provided in Multimedia Appendix 4 [3,6-20,42-45,48-56,58]. The 5 themes are summarized in greater detail in the subsequent sections.

Legal and Institutional Context for Holding and Sharing Data and Its Effect on Ability and Willingness to Share Data

A key challenge to DSCR implementation involved information governance concerns about holding and sharing data. These issues arose from vague legislation, market competition, conflicting priorities, poor internal and external coordination, and low cross-organizational trust. Building trust and adequately resourcing digital change facilitated implementation.

Commercial and Regulatory Context in Which Care Providers Operate

A total of 3 studies conducted in the United Kingdom [3,42,50] identified barriers related to market competition among social care provider organizations and digital suppliers and a lack of national regulation and standards. Private sector care providers were concerned that the commercial sensitivity of data could compromise their competitive advantage [42]. Vendor lock-in also occurred, as technology suppliers hesitated to share data with other suppliers [3].

The governance and ethics framework for social care data in the United Kingdom is less developed compared to that for National Health Service (NHS) data. There is no established system for the governance of care home data, which are held by private companies, care regulators, and health and social care provider organizations [42]. This context made data sharing challenging [42]. Despite regulatory progress, councils found new national frameworks inadequate on data and interoperability standards [50].

Interorganizational Trust and Relationships

In 5 studies [3,9,42,48,49], 4 of which were UK based, a lack of trust between providers and other organizations hampered information governance and data sharing. Clinical and health care partners were particularly reluctant to share data with social care [3,9,50] due to misunderstandings about their role and concerns about sharing information with staff who were not registered social workers [50]. Ownership of a large volume of patient data and responsibility for confidentiality also fostered a risk-averse attitude among general practitioners [3].

Four studies conducted in the United Kingdom [3,42,49,50] found that building trust and collaboration between organizations facilitated implementation. Care homes were more willing to share information when they had well-established relationships with local authorities [42]. Scotland's regional "data safe havens," led by trusted partners such as the NHS, academic

institutions, and government agencies, represented a centralized approach to managing, storing, and handling access requests to health care data that encouraged relationships between health and social care provider organizations [42]. They were an example of data being handled respectfully, professionally, and securely [42]. Ambiguous governance frameworks necessitated clarifying information governance requirements and building mutual trust in systems. In the United Kingdom, local authority and provider staff needed to dedicate significant resources upfront to ensure safe data handling processes [3]. Setting up information sharing agreements that specified data flows between organizations could be intensive, involving unexpected time and effort that was often related to building relationships and engaging numerous actors with data sharing plans [3]. Undertaking this work early on in projects facilitated implementation, locating expertise and capacity, and building trust across organizations [3]. Leaders who fostered positive working relationships between decision makers facilitated shared priority setting [49], helping them to circumvent barriers stemming from organizational fragmentation [49].

Organizational Coordination to Clarify Information Governance

In 4 studies [3,9,42,49], a general lack of coordination hindered the clarification of information governance processes needed to implement digital systems. Divisions between and within organizations created siloed data systems, resulting in residents' records being stored in different systems across multiple services [49]. Poor coordination was linked to information governance professionals, who managed personal data for single organizations, lacking the capacity to handle additional responsibilities for cross-organization information governance and data sharing [3]. This issue was compounded by provider leaders' lack of understanding of information governance [3].

Lack of a shared, standardized understanding of information governance and data ownership across organizations also created confusion among staff. Nursing home leaders in the United States [9], for example, raised concerns about transparency and maintaining control of residents' health data that were viewed as belonging to the patients, leading to fear of lawsuits regarding data sharing [9]. In the United Kingdom, there was a lack of shared understanding with confusion about consent, which was related to social care and local government starting from a different position to NHS partners when it came to information sharing [50].

While a lack of organizational coordination was a barrier to implementation, 4 studies [3,10,49,58] identified that prioritizing and adequately resourcing the work required to define information governance was a facilitator to data sharing and in turn improved service quality. In the implementation of information systems across home care services in Italy [58], agreements could be reached on hardware and software once information governance had been properly defined. The synergies resulting from integrating information systems from different organizations then positively affected service quality. In another study, health and social care managers also acknowledged that undertaking considerable work together to agree on what could be shared helped to implement a shared electronic record between nursing and adult social care practitioners [10]. While fostering cross-organizational relationships was important, substantial resources were required to develop and sustain these relationships [49].

Digital Readiness and Organizational Capacity for Change

The importance of investing in the necessary groundwork and anticipating the work involved in digital implementation is linked to an organization's digital readiness and capacity for new technology more generally. Facilitators and barriers within this theme were related to hardware and internet connectivity issues, funding issues in the sector, organizational infrastructure, and resourcing the work required for digital change, including leadership and training.

Hardware and Internet Connectivity Issues

Hardware issues hindered implementation in 4 studies [6,8,11,55] and negatively impacted care quality in 2 studies [6,11]. Problems included a lack of computers and handheld devices for timely patient data documentation in nursing homes [6,8], ergonomic challenges in home care [55], and poor battery lives on portable devices in home care [11,55]. In Australia, residential aged care nurses and care workers relied on memory when portable devices ran out of battery during medication rounds, reducing patient safety [11]. Sharing limited devices in nursing homes also delayed access to updated care plans in a cross-country nursing home study [6]. Hardware issues implied a failure to commit the upfront investment needed to install the hardware required to successfully implement digital systems, reflecting a lack of organizational capacity and readiness [11,55]. Internet connectivity issues were identified as barriers in 5 studies [13,42,48,55,56]. Reliable internet was often deemed essential for digital implementation, and poor connectivity indicated insufficient organizational resources. This was problematic in home care, where mobile internet access was inconsistent [55,56], and in care homes with poor Wi-Fi in old buildings [13,48]. For instance, 18% of care homes in a southeast Scotland project experienced regular internet interruption, and 27% of care homes had limited internet access [42].

Funding Issues in the Sector

In total, 9 studies [3,9,14,16,18,48-50,56] identified insufficient funding or financial incentives and high costs as barriers. Four studies were conducted in the United States [9,14,16,18], 4 studies were conducted in England [3,48-50], and 1 study was conducted in Sweden [56]. In England, short-term funding pushed organizations toward unambitious digital solutions [3]. The financial pressures often forced providers to adopt a short-term view on the finances needed to implement and sustain digital records, constraining the scale of change and preventing it from being embedded [3]. Where funding was available, finding, requesting, and receiving it was not always straightforward [48]. Small care homes in England faced issues such as poor communication from funders, complicated application procedures, and delays in receiving funds [48].

Organizational Infrastructure and Resourcing the Work Required for Digital Change

Barriers related to organizational infrastructure were noted in 5 studies [19,20,48,49,56]. Issues included insufficient ICT and human resources staff [48,56], high senior staff turnover [49], poor internal communication that left staff unaware of implementation [19,20], and inadequate leadership [20]. One English study [3] highlighted that successful implementation required clear planning and resource allocation; for example, phased deployment of resources demonstrated providers' competence in managing digital change, making it easier for them to secure further funding [3].

Four studies [3,10,50,53], 3 of which were conducted in England, noted the importance of skillful leadership in enhancing digital readiness. Identifying leaders with the right skills was crucial for managing large-scale digital projects [3]. The type of leadership required depended on context, with some providers preferring leaders who could balance risk and reward in deploying resources, while others sought leaders who were respected by their peers to help foster engagement among staff [3]. Senior staff functioning as "change agents" also motivated practitioners to review their practices [10]. Successful councils had strong leadership support for digital initiatives [50]. In England, councils successful in implementing data standards and interoperability had strong leadership support [50], with directors of social care, chief information officers, and elected members all prioritizing digital working and integrated care [50].

Adequate Training

Absent or inadequate training was a barrier in 4 studies [6,19,20,48]. Issues included a lack of tailored training [20] and inappropriate content [19]. Conversely, 5 studies identified high-quality training as a facilitator [3,6,11,49,53]. One multicountry study identified both facilitators and barriers across the different contexts [6]. Effective training was tailored to practitioners' skills and tasks [6,53] and included on-the-job and context-specific training [6], ongoing sessions [55], follow-up visits [53], and continued onsite support from suppliers [6,11]. High-quality training that was tailored, targeted, and practical aligned care practices with the new practices required by digital systems.

Using and Sharing Recorded Information Within Technical Constraints

This theme included issues with technical interoperability of digital systems, their level of usability and user-friendliness, and the extent to which they had been appropriately adapted for social care from other settings, which were often acute or primary care.

Interoperability

Interoperability is understood as a technology's capacity to electronically share patient information between different systems and to use the information that has been shared [59]. Lack of interoperability was identified in 10 studies as a barrier to sharing recorded information [3,6,9,15,17,42,49,50,52,54], being reported by 57% of 491 respondents in the US [15] nursing facilities with electronic health records. Care professionals and managers in Finland [52] and senior health and care leaders in England [49] also criticized information systems for not always "communicating" with each other. While providers were adopting digital solutions, these were not necessarily increasing interoperability and risked creating new data silos [15].

In some studies, interoperability barriers were attributed to the multitude of systems used by different organizations. Across 9 nursing homes in Austria [17], managers exchanged information with at least 18 other organizations, most of which were not part of the same electronic health record system. In the United States, while 95.1% (775/815) of nursing homes had electronic medical records, only 45.8% (373/815) had some capability for information exchange with other organizations. The variety and sheer number of systems used by different providers was a concern for 8 (67%) out of 12 staff members in subsequent interviews [9].

In England [50], interoperability issues presented as systems being unable to store identification data such as the NHS number. However, local authorities were often unaware, at the procurement stage, of which digital options could store such information. There was also confusion among councils and suppliers about the possibilities and limitations of NHS number tracing. This was linked to low organizational readiness and capacity, with providers not knowing which technological features they needed when choosing a system [50]. It also related to the downstream value suppliers promised providers in terms of being transparent about what their products could offer [50].

Staff in all 3 nursing homes in a multicountry study also complained that the electronic patient records lacked interoperability and options to adjust features to meet specific needs. This implied a contradiction between customizability and interoperability, with customizable systems more likely to meet care provider needs but less likely to be compatible with other systems than off-the-shelf technology [6].

Usability and User-Friendliness

A total of 8 studies [6,11,12,51,52,54-56] reported barriers related to this theme, 3 of which [52,54,56] were based in Nordic countries. These barriers were more closely associated with using, rather than sharing, recorded information within technical constraints.

A total of 4 studies [51,52,54,55] found problems with the system being slow, crashing, and having unscheduled downtime. Others pointed to features that made staff work routines more inefficient, such as the example from US home care nurses needing to click 22 times to get into each individual's medical record, a cumbersome process that had to start again when they moved on to the next patient [12]. In a multicountry study, care home staff disliked being forced to enter narrative text into the electronic patient record and preferred drop-down menus [6]. An inefficient information retrieval process within an Australian electronic Health record system meant that staff in long-term care facilities had to perform lengthy manual searches to identify wound charts, with the system also failing to alert them if they were duplicating charts that already existed [11].

In 5 studies [8,11,13,20,55], the high usability and usefulness of digital systems facilitated implementation. In 3 cases [13,20,55], systems offered easy access to information, improving the immediacy of care provision and documentation. In some instances, they enhanced the accuracy of care documentation through better information visibility [13,55] or by automating tasks that were previously manual and prone to human error [11]. Digital systems with flagging features also supported resource prioritization and management decisions [13]. These facilitators aided implementation by increasing task efficiency and supporting the knowledge generated or made visible by the technology, thereby improving data accuracy and decision-making.

Adapting Technology From Other Settings

Barriers in 5 studies [6,20,45,54,56] were related to digital systems that had been maladapted from other settings and were consequently deemed inappropriate for social care. In England, social care workers were less likely to perceive health information exchange systems as useful compared to health care workers and experienced issues with the user interface [20]. Staff noted that the system looked unfamiliar compared to other systems they used, as the health information exchange was primarily designed for acute and primary care settings, with little consideration given to social and community services [20].

An Australian study [13] reported successful adaptation of a digital system originally designed for an acute hospital setting to a care home involving staff at all levels, residents, and their relatives that helped to make the product appropriate for the care home setting [13]. This co-designed process facilitated implementation and increased the likelihood of success.

Alignment Between Care Practices and Digital Recording Practices

Overview

Barriers related to digital systems not matching the context of use were identified in 9 studies [3,6,7,12,45,50,54-56]. These barriers referred to misalignments between care practices within the social care sector and recording practices demanded by new digital systems. They included reduced interactions between clients and practitioners, conflicts with preferred data input methods, and exacerbation of existing organizational issues. Staff perceptions of improved care quality increased the likelihood of accepting the technology.

Care Quality and the Relational Nature of Social Care

A total of 5 studies [6,7,9,49,52] highlighted barriers where digital systems decreased the relational nature of social care. Problems arose when care staff experienced disruption to their relational work and viewed the technology as depersonalizing care. In Finland, new information systems increased technical tasks at the expense of relational tasks performed physically close to clients [52]. In England, digital records influenced the nature of the clinical encounter for occupational therapists. By focusing on data collection and adherence to standard procedures, they reduced opportunities for building rapport with clients [7]. Concerns also existed that technology use close to clients was intrusive and reduced care quality [6,55]. Defining the problem as a preference for “high touch” over “high tech,” a US study found that 5 out of 12 nursing home leaders feared that technology might detract from the personal experience they aimed to provide [9].

Technological features, such as prescriptive data fields, also imposed work routines that prioritized clinical data and processes. In a Scottish study, data systems in care homes promoted a task-oriented culture over resident-focused care [45]. Prescripted data fields limited the recording of social and emotional activities and care provision, leading to an overly clinical focus in the data [45].

Only 1 study found that a digital system aligned well with the relational nature of social work, facilitating implementation [13]. In Australia, an aged care ecosystem that was co-designed with staff and residents allowed care workers to multitask and spend more time with residents. This saved time for staff and improved care quality, encouraging acceptance of the system [13]. Managers noted that the technology provided prompts for tasks such as repositioning residents, better aligning care with resident needs [13]. In England, 2 studies found that perceived care quality improvements increased staff acceptance of digital systems [3,20]. Demonstrating the technology’s value to different professionals helped staff “buy into” digital change [3]. Administrative staff adopted technology for time-saving benefits, while practitioners focused on its impact on care [3]. Perceived improvements to patient safety also increased the likelihood of adopting digital systems [20].

Preexisting Organizational Problems

A total of 2 studies [12,19] identified barriers where digital systems exacerbated preexisting organizational problems, such as the numerous communication channels in home care organizations [12]. The lack of standardization required nurses to adapt to various communication methods, for example, contacting physicians through primary care nurses or by fax [12]. They often only discovered that their request had reached doctors through changes made to patients’ medicines [12]. Rather than standardizing processes, the new digital system added more communication channels. While this issue presented as inappropriate technology, it was rooted in inefficient work routines that predated the technology’s introduction.

Conflicts Between Data Recording Practices and Digital Systems

A total of 4 UK-based studies [3,7,42,50] identified barriers due to conflicts between data recording practices preferred by care providers and those permitted by digital systems. The lack of systematic data collection in care homes made it difficult to capture the complexity of care for individuals with multiple conditions and high support needs [42]. Frontline practitioners preferred narrative text input, while digital systems often emphasized coded data entry [3,50]. In 1 study, social workers entering free-text information sometimes included data about third parties without consent [3]. Such issues were linked to a lack of understanding about data quality in social care [50], requiring retraining on the importance of proper data collection and recording practices [3,50]. An English study found that conflicts between recording preferences and the recording permitted by digital systems were due to a mismatch between digital care records and occupational therapy concerns [7].

Therapists had to recode their interventions to fit the system's structure, suggesting that the technology did not align with sector needs, rather than indicating poor recording practices.

Differences Between What Is Expected and What Is Achievable With Digital Systems

Overview

The final theme related to the gap between organizational expectations and realistic achievements with digital systems. Guidance on available technology was often inadequate, and care providers lacked internal consensus about the technology's capabilities and what they wanted to gain from implementation. Creating a shared digital vision and adopting digital systems as part of wider cultural changes facilitated implementation.

Guidance on the Technology Available

Insufficient guidance on available technology was a barrier in 1 study in English care homes [48]. The overwhelming number of suppliers created an "unregulated tech product maze," making it difficult to choose the best option and avoid paying for unsuitable technology [48]. Care homes criticized NHS England's "assured suppliers list" of DSCR suppliers, which was introduced to aid decision-making [48]. Although suppliers on the list met a set of standards, some care homes complained that suppliers did not meet their needs and requirements, while others reported poor experiences with suppliers on the list and found themselves locked into contracts despite consistent software malfunctions [48].

A Shared Digital Vision

Creating a shared vision for collectively understanding the technology involved building organizational consensus on its potential while remaining realistic about its limitations. A total of 4 studies [3,17,20,52] found that care provider staff disagreed about the purpose of digital systems, and awareness of potential benefits for care delivery was low. There were tensions between 2 distinct staff groups with different expectations [3]. One group represented a technical and managerial culture that often initiated digital change projects and was primarily interested in the information captured by digital systems. The other was a clinical culture that was concerned with how technology could help deliver care and was more skeptical of changes to practice that lacked certain types of evidence [3]. Managers were generally more positive about implementation but lacked awareness of some of its negative effects on employees' work [52]. Staff anticipated unrealistic benefits and were often unaware of the technology's value [17,20].

A total of 2 studies found facilitators to creating a shared vision [13,53]. They highlighted the importance of co-design and inclusive implementation by gathering suggestions from staff, residents, and their relatives [13] or by conducting monitoring based on staff's feedback to system developers [53]. Involving different groups as partners in the process helped envision a digital system that benefited everyone [13]. Professionals praised comprehensive and continuous communication that helped them make sense of a new service, with information delivered through multiple channels to reach as many employees as possible, including shift workers [53].

Implementing Digital Change as a Cultural Change

Framing digital implementation as a cultural change program facilitated success in 3 studies [3,10,13]. In an Australian care home, co-designing the system, establishing a shared vision across the workforce, and providing training and feedback loops instigated a culture change that improved service delivery and problem-solving [13]. In England, barriers to scaling digital changes in health and social care were mitigated by treating them as part of a wider technology-supported clinical transformation program, rather than an ICT project [3], or as part of a larger cultural change program to improve administrative efficiency [10].

Summary of the Intended and Actual Outcomes

Outcomes of digital implementation, either intended or actual, were identified in 17 studies [7-11,13,17,19,20,48,50-52,54,55,57,58], although they were the focus of only 1 study [57]. The full details of the benefits and outcomes are provided in Table 2.

A total of 3 studies [9,17,48] identified the outcomes that participants hoped to achieve through adopting digital systems. Improved information accessibility, information sharing, and quality of care records were identified in 2 studies [17,48], making them the most frequent intended outcomes. Examples of the improved quality of records included more complete and readily available patient-related information and less documents being lost during patient transitions between different institutions [17]. Improved efficiency [17] and time savings [48] were identified as intended outcomes in 1 study.

Three studies [9,48,50], 2 of which were based in England [48,50], cited poor awareness about the benefits of digital systems for social care or concerns that they would not benefit the sector. In England, information sharing initiatives were often focused on health care and hospitals, with less attention paid to the potential benefits for councils or social care [50]. This made it difficult for social care staff and care home residents to see the benefits that digital systems could bring [48].

A total of 13 studies identified positive outcomes realized through digital record implementation [7,8,10,11,13,19,20,50-52,55,57,58]. Improved efficiency was the most frequent actual outcome (8/13, 62%) [10,11,13,19,20,50-52], achieved through the automation of previously manual processes [11], reduced duplication of procedures [20], and the increased availability [19] and immediacy [13] of information improving decision-making and care planning. These outcomes were associated with increased staff capacity [50] and productivity [52]. Impacts on efficiency were not always clear. In 1 study [11], while automatic data entry in patient records was beneficial, the system did not completely align with work processes, and staff needed to record some data twice.

Table 2. The intended and actual outcomes of digital social care record implementation (N=17).

Theme	Studies, n (%)
Intended outcomes	
Improved quality of data records	2 (12)
Improved information sharing	2 (12)
Improved information accessibility	2 (12)
Improved efficiency	1 (6)
Time savings	1 (6)
Improved care quality or planning	1 (6)
Improved communication or collaboration	1 (6)
Improved information accuracy	1 (6)
Space savings (less paper)	1 (6)
Actual outcomes	
Improved efficiency	8 (47)
Perceived time savings	7 (41)
Improved information accessibility	5 (29)
Workarounds (viewed negatively)	4 (24)
Improved communication or collaboration	3 (18)
Improved information security and risk management	3 (18)
Additional time burdens	3 (18)
Improved care quality or planning	2 (12)
Increased face-to-face work with patients	2 (12)
Improved information sharing	2 (12)
Improved information accuracy	2 (12)
Improved transparency and accountability	2 (12)
Increased staff or patient satisfaction	1 (6)
Workarounds (viewed positively)	1 (6)
Decreased communication or collaboration	1 (6)
Decreased efficiency	1 (6)
Decreased care quality	1 (6)
Decreased face-to-face work with patients	1 (6)
Lack of financial benefits	1 (6)
Rationing care documentation	1 (6)

Perceived time savings were reported in 7 studies, although the findings were not conclusive [10,13,19,20,50,52,58]. Some studies reported staff spending less time retrieving and documenting information for decision-making [13,58] and chasing other organizations for patients' whereabouts [50]. One study found time savings of up to 45 minutes for long-term care staff when completing medication reconciliation [19]. However, 2 studies found time savings in some areas and additional time burdens in others [10,52]. In 1 case, disagreements between managers and their staff arose regarding whether the digital system created time savings [52]. Managers and employees agreed that moving from phone calls to digital messaging had freed up staff time for other tasks [52]. However, employees felt that the new tasks, such as responding to clients through messages, required extra time. This additional time was not always recognized by management, nor were additional resources provided [52].

A total of 3 studies [8,13,19,20,52] found that digital systems made information more accessible. In one case, this enabled person-centered care, with easily accessible information on individual backgrounds helping staff to "see the person first and the diagnosis second" [13]. In another case, improved visibility of information facilitated medication tracking and therefore supported patient safety [19]. A total of 3 studies also highlighted improved communication and collaboration [10,52,55] and improved information security and risk management [51,52,58]. Electronic information sharing

improved partnership working, enhancing collaboration and increasing the timeliness, efficiency, and quality of care [10,52,55]. Improved information security and risk management were linked to secure information transfer and storage [51,52,58], better client monitoring [52], and increased data accuracy [51,58].

Workarounds, identified in 5 studies [7,9,11,54,55], were the most common negative outcome. Workarounds involve the implementation, by end users, of temporary practices or behaviors to overcome the limitations of a technological system [60]. Staff developed workarounds for various reasons. These included circumventing the system to share health data with residents [9] and accessing case-based information [54]. While workarounds could be beneficial [7] and support task completion [54], they also threatened data security [54].

Discussion

Principal Findings

This study investigated what is known about the implementation of digital records in adult social care settings. The literature was diverse in terms of the type of digital system, setting, and use case studied. Most of the studies used a qualitative design (14/29, 48%), particularly those looking at facilitators and barriers, how digital systems affect work routines, and potential strategies to improve digital systems. Studies were most frequently based within the United Kingdom (10/29, 34%).

Most studies focused on facilitators of and barriers to digital implementation. Many facilitators and barriers were interlinked and associated with multiple NASSS framework domains, which compounded the complexity of implementing digital systems. The 5 themes we identified using the NASSS framework are particularly complex areas that require more active management and consideration when implementing DSCRs in social care contexts.

While our findings suggest that implementing digital systems is an inherently complex process, this review did identify some strategies to manage complexity, which could constitute "good practice." In terms of digital readiness and organizational capacity for change, high-quality training was found to increase implementation success. Where training was tailored, practical, and ongoing, it helped align care practices with new practices required by the technology, thereby increasing employees' ability and willingness to adopt and continue to use the system. Although high-quality training depended on care provider leaders anticipating the financial resources needed, it seems a worthwhile investment for successful digital implementation. This finding echoes the results from a previous scoping review, which highlighted training as a key factor influencing the use of electronic information systems [23].

Implementing digital systems as part of wider cultural change projects also addressed multifaceted complexity. An example of this was the project in which implementation was co-designed with staff [13]. This approach enabled a shared vision of the technology to be created across the care home among residents and staff at different levels. The sense of ownership this instilled addressed complexity in the adopter system domain, with all users more likely to support the technology and view it as "business as usual." Co-design also addressed complexity in the technology domain, with the digital system more likely to align with the needs and practices of its user group. While incorporating digital implementation as part of broader transformation required significant resources, where there was sufficient organizational readiness and capacity for comprehensive rollout, implementation seemed to have greater potential for sustainability, scaling, and spread.

Complexity related to data sharing and information governance seemed to be more difficult to address. Trust and relationship building across organizations could help establish data sharing agreements at a localized level and therefore address complexity within the organization domain. However, fundamental barriers were associated with complexity around regulations and standards in the wider context domain, over which care providers had no direct control. Until there is primary or secondary legislative change, the governance and regulatory context will continue to impede cross-organizational data sharing efforts.

Although 17 studies identified intended or actual outcomes, they more often focused on identifying facilitators of and barriers to implementation. Improved efficiency, accuracy, and time savings were the most common positive outcomes realized through digital adoption, while workarounds and additional time burdens were the most frequently cited negative outcomes. Some of the positive outcomes reflect the results presented in the review by Greenstock [22], which also found efficiency and productivity to be a benefit of DSCRs. However, the limited detail in outcome reporting and variations in the extent to which different benefits are observed suggest that this topic would benefit from future research. Specifically, there seems to be a need for studies that quantify outcomes and pay greater attention to the necessary conditions for positive benefits to be realized.

Limitations of Studies

Most studies (20/29, 69%) lacked a clear theoretical or methodological framework. This meant it was often unclear which type of digital system or record was being implemented as well as the context, setting, and use case. While studies mentioned >100 facilitators and barriers to implementation, they did not provide any objective parameters or measures to assess how they impact implementation or social care practices. This hinders a more comprehensive comparison between the barriers and facilitators.

Some digital systems were simply described as ICT, electronic digital systems or health information technology [18,55,57], or digitalization or digital change generally [3,48,52], without definitions of these terms. Some studies appeared to use the same vocabulary to describe different systems. However, this was difficult to determine as most studies (20/29, 69%) did not specify their target population clearly. Many studies also lacked detail regarding care settings and other relevant information, which limited the possibility of performing more comprehensive comparative analysis. Future studies should pay greater attention to how they report which digital systems were implemented, the target population for the system, the setting, and the roles of the professionals involved to facilitate comparisons between

studies. Standardized reporting guidelines, such as the template for intervention description and replication checklist and guide [61], may facilitate describing digital projects or systems.

Of the 29 studies, only 2 (7%) included carers or service users as respondents, while most studies included staff (n=18, 62%) or managers (n=11, 38%). Future research may therefore benefit from incorporating the perspectives of people drawing on care to cover this gap in the literature.

Although studies included in our review mention the potential impacts of DSCRs, none provide quantifiable parameters to estimate such impacts, such as potential time savings or cost-effectiveness metrics. New studies that are appropriately designed to measure such outcomes are needed to fill this important knowledge gap in the literature on DSCR implementation.

Methodological Limitations

Due to the prevalence of qualitative designs and a lack of clear theoretical or methodological frameworks among the studies reviewed, we used the NASSS framework as a structured approach to categorizing and interpreting heterogeneous data. As this was a rapid review, the framework served as a tool to guide our data interpretation and triangulation, especially given the large number of barriers and facilitators and the varied ways these issues were described across different studies. For example, it directed our analysis of hardware and internet connectivity issues. While the studies reviewed often attributed these to technological problems, the framework enabled us to trace the associated complexity back to the organization domain, with care providers lacking the awareness, readiness, and capacity to prepare for digital implementation and adopt appropriate systems.

However, the NASSS framework carried some limitations for our analysis. No facilitators or barriers were associated with the condition domain. While the framework was developed for both health and social care, the focus of this domain on comorbidities and clinical aspects of a patient's condition may be more appropriate for health care technologies. For social care technologies, it may be more useful to approach the condition domain in terms of whether digital systems are appropriate for particular groups of clients, such as older people or people with learning disabilities, rather than specific illnesses. Alternatively, the lack of relevance of the condition domain may reflect limited attention to diversity and inclusion considerations within the studies reviewed. Another limitation of the NASSS framework was related to the final domain (ie, interaction between domains and adaptation over time). As most of the complexity we identified was multifaceted, we found it more useful to iteratively analyze the interactions between domains instead of restricting them to 1 domain. Rather than viewing complexity as belonging to separate domains, we suggest using this final domain to provide an overarching perspective of how complexity constantly intersects and interacts across domains at every stage of digital implementation.

Considering the rapid nature of this scoping review, we simplified some review procedures, such as screening and full-text assessment, which always carry the risk of missing relevant studies. To minimize such risks, all review procedures were undertaken by researchers who are experts in systematic review methods and social care research. As with every review, the choice of databases is also a limitation, as relevant studies may have been uniquely indexed in databases that were not included. However, our research was able to identify all relevant studies that were suggested by experts in the field. Moreover, we performed a comprehensive gray literature search to reduce the likelihood of missing key studies.

Despite the limitations, we believe that our review provides a comprehensive picture of the state of the literature on DSCRs. It builds on 4 previous reviews, which, when taken together, captured the literature about digital records until the end of 2017 [2,21-23]. Our review has updated and added to these findings, covering both academic and gray literature up until 2023 and using a robust theoretical framework to draw out complexity in terms of sustainability, scaling, spread, nonadoption, and abandonment of digital care records.

Conclusions

Our findings suggest that the implementation of digital care records is particularly complex due to the lack of a common language and consensus about what DSCRs should look like as well as expected outcomes and impacts. This is reflected in the scientific literature, which often lacks operationalization of key constructs and robust study designs. To be successful, implementation should consider complexity, while studies should use a robust theoretical framework and use mixed methods or quantitative designs where appropriate. We also suggest that future studies define the target population, consider gathering data on the experiences of carers and service users, and focus on digital care records specifically being used in social care, such as DSCRs.

Acknowledgments

The authors would like to acknowledge the support of the Social Care Rapid Evaluation Team coordinator Jo Ward, as well as Andra Fry, the London School of Economics and Political Science library information specialist. This study is funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research Programme (NIHR163639) as part of the Social Care Rapid Evaluation Team (NIHR153673), and open access fees were covered by the grant. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Data Availability

No new data were generated for this scoping review. All data were obtained from publicly available published sources and are included in this published article and its supplementary information files. We have included the search strategy in Multimedia Appendix 2, summaries of the extracted articles in Multimedia Appendix 3, and the thematic analysis process in Multimedia Appendix 4 to provide transparency on extractions and analysis.

Authors' Contributions

MS contributed to screening and data collection and was responsible for data analysis and preparation of the first draft of the manuscript. WS-R contributed to study conception and design, was responsible for screening and data collection, and contributed to data analysis and preparation of some sections of the manuscript. JM conceived and designed the research; oversaw the screening, data collection, and data analysis; and contributed to drafting of the manuscript. KP and SDG were responsible for developing and running the search strategy and drafting the paragraphs on the search methods. MB and CL contributed to the study conception and design. WS-R, JM, MB, JW, CL, NF, and NS critically reviewed the manuscript, adding important intellectual content. All authors read and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 162 KB - aging_v8i1e60107_app1.pdf]

Multimedia Appendix 2
Search Strategy, August 14, 2023.
[DOCX File , 32 KB - aging_v8i1e60107_app2.docx]

Multimedia Appendix 3
Detailed characteristics of the included studies (N=29).
[DOCX File , 43 KB - aging_v8i1e60107_app3.docx]

Multimedia Appendix 4
Thematic analysis of the facilitators and barriers using the nonadoption, abandonment, scale-up, spread, and sustainability framework.
[DOC File , 64 KB - aging_v8i1e60107_app4.doc]

References

1. Bongaarts J. Human population growth and the demographic transition. Philos Trans R Soc Lond B Biol Sci 2009 Oct 27;364(1532):2985-2990 [FREE Full text] [doi: 10.1098/rstb.2009.0137] [Medline: 19770150]
2. Kruse CS, Mileski M, Vijaykumar AG, Viswanathan SV, Suskandla U, Chidambaram Y. Impact of electronic health records on long-term care facilities: systematic review. JMIR Med Inform 2017 Sep 29;5(3):e35 [FREE Full text] [doi: 10.2196/medinform.7958] [Medline: 28963091]
3. Maguire D, Evans H, Honeyman M, Omojomolo D. Digital change in health and social care. The King's Fund. 2018 Jun 20. URL: https://www.kingsfund.org.uk/insight-and-analysis/reports/digital-change-health-social-care [accessed 2025-02-05]
4. Health at a glance 2023. Organisation for Economic Co-operation and Development. 2023 Nov 7. URL: https://www.oecd.org/en/publications/2023/11/health-at-a-glance-2023_e04f8239.html [accessed 2025-02-05]
5. Slawomirski L, Lindner L, de Bienassis K, Haywood P, Hashiguchi TC, Steentjes M, et al. Progress on implementing and using electronic health record systems: developments in OECD countries as of 2021. Organisation for Economic Co-operation and Development. 2023 Sep 13. URL: https://tinyurl.com/4fspm6ry [accessed 2025-02-05]
6. Shiells K, Diaz Baquero AA, Štěpánková O, Holmerová I. Staff perspectives on the usability of electronic patient records for planning and delivering dementia care in nursing homes: a multiple case study. BMC Med Inform Decis Mak 2020 Jul 13;20(1):159 [FREE Full text] [doi: 10.1186/s12911-020-01160-8] [Medline: 32660474]
7. Sugarhood P, Rouncefield M. Occupational therapists and digital care records: exploring health and social care integration through ethnography - summary of key findings. UK Occupational Researcher Foundation. 2018. URL: https://www.rcot.co.uk/sites/default/files/UKOTRF%20Key%20findings_Dr%20Paul%20Sugarhood.pdf [accessed 2025-02-05]
8. Ausserhofer D, Favez L, Simon M, Zúñiga F. Electronic health record use in Swiss nursing homes and its association with implicit rationing of nursing care documentation: multicenter cross-sectional survey study. JMIR Med Inform 2021 Mar 02;9(3):e22974 [FREE Full text] [doi: 10.2196/22974] [Medline: 33650983]

9. Powell KR, Deroche CB, Alexander GL. Health data sharing in US nursing homes: a mixed methods study. J Am Med Dir Assoc 2021 May;22(5):1052-1059 [FREE Full text] [doi: 10.1016/j.jamda.2020.02.009] [Medline: 32224261]
10. Chester H, Hughes J, Bowns I, Abendstern M, Davies S, Challis D. Electronic information sharing between nursing and adult social care practitioners in separate locations: a mixed-methods case study. J Long Term Care 2021 Jan 12:1-11 [FREE Full text] [doi: 10.31389/jltc.16]
11. Qian S, Yu P, Bhattacherjee A. Contradictions in information technology mediated work in long-term care: an activity theoretic ethnographic study. Int J Nurs Stud 2019 Oct;98:9-18. [doi: 10.1016/j.ijnurstu.2019.05.017] [Medline: 31238234]
12. Persson J, Larsson R, Erlingsdottir G, Rydenfält C. How digital systems are used in Swedish home care nursing practice: a qualitative interview study to identify challenges and opportunities. Comput Inform Nurs 2023 Aug 01;41(8):586-594 [FREE Full text] [doi: 10.1097/CIN.0000000000001006] [Medline: 36649078]
13. Bail K, Gibson D, Hind A, Strickland K, Paterson C, Merrick E, et al. 'It enables the carers to see the person first': qualitative evaluation of point-of-care digital management system in residential aged care. J Clin Nurs 2023 Jan;32(1-2):174-190 [FREE Full text] [doi: 10.1111/jocn.16285] [Medline: 35285557]
14. Felix H, Dayama N, Morris ME, Pradhan R, Bradway C. Organizational characteristics and the adoption of electronic health records among nursing homes in one southern state. J Appl Gerontol 2021 May 21;40(5):481-488. [doi: 10.1177/0733464820906685] [Medline: 32081058]
15. Vest JR, Jung HY, Wiley KJ, Kooreman H, Pettit L, Unruh MA. Adoption of health information technology among US nursing facilities. J Am Med Dir Assoc 2019 Aug;20(8):995-1000.e4 [FREE Full text] [doi: 10.1016/j.jamda.2018.11.002] [Medline: 30579920]
16. Lin SC, Tunalilar O. Rapid adoption of electronic health record and health information exchange among assisted living communities, 2010-2018. J Am Med Inform Assoc 2022 Apr 13;29(5):953-957 [FREE Full text] [doi: 10.1093/jamia/ocac021] [Medline: 35187569]
17. Schaller M, Dornauer V, Hackl WO, Lechleitner G, Überegger M, Ammenwerth E. Implementing national electronic health records in nursing homes in Tyrol: a nursing management perspective. Stud Health Technol Inform 2020 Jun 23;271:240-247. [doi: 10.3233/SHTI200102] [Medline: 32578569]
18. Alexander GL, Liu J, Powell KR, Stone PW. Examining structural disparities in US nursing homes: national survey of health information technology maturity. JMIR Aging 2022 Aug 23;5(3):e37482 [FREE Full text] [doi: 10.2196/37482] [Medline: 35998030]
19. Aird T, Holditch C, Culgin S, Vanderheyden M, Rutledge G, Encinareal C, et al. An analysis of a novel Canadian pilot health information exchange to improve transitions between hospital and long-term care/skilled nursing facility. J Integr Care 2022 Sep 19;30(4):399-412. [doi: 10.1108/jica-03-2022-0022]
20. Watkinson F, Dharmayat KI, Mastellos N. A mixed-method service evaluation of health information exchange in England: technology acceptance and barriers and facilitators to adoption. BMC Health Serv Res 2021 Jul 25;21(1):737 [FREE Full text] [doi: 10.1186/s12913-021-06771-z] [Medline: 34303379]
21. Kruse CS, Marquez G, Nelson D, Palomares O. The use of health information exchange to augment patient handoff in long-term care: a systematic review. Appl Clin Inform 2018 Oct;9(4):752-771 [FREE Full text] [doi: 10.1055/s-0038-1670651] [Medline: 30282094]
22. Greenstock J. Benefits of digital social care records. The Strategy Unit. 2021 Aug. URL: https://tinyurl.com/3p87s34y [accessed 2025-02-05]
23. Ylönen K. The use of Electronic Information Systems in social work. A scoping review of the empirical articles published between 2000 and 2019. Eur J Soc Work 2022 Apr 18;26(3):575-588. [doi: 10.1080/13691457.2022.2064433]
24. Peek N, Sujan M, Scott P. Digital health and care: emerging from pandemic times. BMJ Health Care Inform 2023 Oct;30(1):e100861 [FREE Full text] [doi: 10.1136/bmjhci-2023-100861] [Medline: 37832967]
25. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
26. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
27. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
28. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
29. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017. URL: https://iris.who.int/bitstream/handle/10665/258698/9789241512763-eng.pdf [accessed 2025-02-05]
30. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]

31. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39 [FREE Full text] [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
32. Ribeiro WS, Malley J, Baginsky M, Larkins C. Good practice in the implementation of digital social care records (DSCRs): a rapid scoping review. Open Science Framework Home. 2023 Aug 9. URL: https://osf.io/d5pcj/ [accessed 2024-04-25]
33. Falconer J. Removing duplicates from an EndNote library. Library, Archives & Open Research Services at the London School of Hygiene & Tropical Medicine. 2018 Dec 7. URL: https://blogs.lshtm.ac.uk/library/2018/12/07/removing-duplicates-from-an-endnote-library/ [accessed 2025-02-05]
34. Covidence systematic review software. Veritas Health Innovation. URL: https://www.covidence.org [accessed 2025-01-22]
35. Local Government Association homepage. Local Government Association. URL: https://www.local.gov.uk [accessed 2023-08-11]
36. The King's Fund homepage. The King's Fund. URL: https://www.kingsfund.org.uk [accessed 2023-08-11]
37. The Social Care Institute for Excellence homepage. The Social Care Institute for Excellence. URL: https://www.scie.org.uk [accessed 2023-08-11]
38. The Centre for Care homepage. The Centre for Care. URL: https://centreforcare.ac.uk [accessed 2023-08-11]
39. Digital Care Hub homepage. Digital Care Hub. URL: https://www.digitalcarehub.co.uk [accessed 2023-08-11]
40. TEC Services Association homepage. TEC Services Association. URL: https://www.tsa-voice.org.uk [accessed 2023-08-11]
41. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC Med 2015 Sep 16;13:224 [FREE Full text] [doi: 10.1186/s12916-015-0465-6] [Medline: 26377409]
42. Johnston L, Koikkalainen H, Anderson L, Lapok P, Lawson A, Shenkin SD. Foundation level barriers to the widespread adoption of digital solutions by care homes: insights from three Scottish studies. Int J Environ Res Public Health 2022 Jun 16;19(12):7407 [FREE Full text] [doi: 10.3390/ijerph19127407] [Medline: 35742667]
43. Johnston L, Kokkalainen H, Anderson L, Lapok P, Lowson A, Shenkin S. Digital and data readiness of care homes for older people in South East Scotland. Edinburghdinburgh Napier University. 2022 Apr. URL: https://edwebcontent.ed.ac.uk/sites/default/files/atoms/files/ladder_study_2022.pdf [accessed 2025-02-05]
44. Johnston L, Hockley J, Henderson DA, Shenkin SD. The development of a care home data platform in Scotland: insights from the care home innovation partnership, Lothian. medRxiv Preprint posted online on August 20, 2020 [FREE Full text] [doi: 10.1101/2020.08.17.20176503]
45. Shenkin SD, Johnston L, Hockley J, Henderson DA. Developing a care home data platform in Scotland: a mixed methods study of data routinely collected in care homes. Age Ageing 2022 Dec 05;51(12):afac265 [FREE Full text] [doi: 10.1093/ageing/afac265] [Medline: 36580390]
46. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
47. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
48. Supporting small care home providers and their managers on their journey towards digitalisation. National Institute for Health and Care Research. 2022 Nov 30. URL: https://tinyurl.com/bddvrtmf [accessed 2025-02-05]
49. Ingram E, Cooper S, Beardon S, Körner K, McDonald HI, Hogarth S, et al. Barriers and facilitators of use of analytics for strategic health and care decision-making: a qualitative study of senior health and care leaders' perspectives. BMJ Open 2022 Feb 17;12(2):e055504 [FREE Full text] [doi: 10.1136/bmjopen-2021-055504] [Medline: 35177457]
50. Local government social care data standards and interoperability. Local Government Association. 2019 Sep 20. URL: https://www.local.gov.uk/publications/local-government-social-care-data-standards-and-interoperability [accessed 2025-02-05]
51. Gatawa T, Swift N, Gibson S, Lindsay-Walters F. Improving social care through digital care planning: an evaluation of the PASSsystem. Social Care Institute for Excellence and everyLIFE Technologies. 2019 Mar. URL: https://everylifetechnologies.com/wp-content/uploads/2020/11/everyLIFE-SCIE-full-report-compressed.pdf [accessed 2025-02-05]
52. Kaihlanen AM, Laukka E, Nadav J, Närvänen J, Saukkonen P, Koivisto J, et al. The effects of digitalisation on health and social care work: a qualitative descriptive study of the perceptions of professionals and managers. BMC Health Serv Res 2023 Jun 30;23(1):714 [FREE Full text] [doi: 10.1186/s12913-023-09730-y] [Medline: 37386423]
53. Nadav J, Kaihlanen AM, Kujala S, Laukka E, Hilama P, Koivisto J, et al. How to implement digital services in a way that they integrate into routine work: qualitative interview study among health and social care professionals. J Med Internet Res 2021 Dec 01;23(12):e31668 [FREE Full text] [doi: 10.2196/31668] [Medline: 34855610]
54. Salovaara S, Ylönen K. Client information systems' support for case-based social work: experiences of Finnish social workers. Nordic Soc Work Res 2021 Nov 10;12(3):364-378. [doi: 10.1080/2156857x.2021.1999847]
55. Ibrahim S, Donelle L, Regan S, Sidani S. A qualitative content analysis of nurses' comfort and employment of workarounds with electronic documentation systems in home care practice. Can J Nurs Res 2020 Mar 14;52(1):31-44. [doi: 10.1177/0844562119855509] [Medline: 31200603]
56. Rydenfält C, Persson J, Erlingsdottir G, Johansson G. eHealth services in the near and distant future in Swedish home care nursing. Comput Inform Nurs 2019 Jul;37(7):366-372. [doi: 10.1097/CIN.0000000000000536] [Medline: 31135467]

57. Yoshimoto T, Nawa N, Uemura M, Sakano T, Fujiwara T. The impact of interprofessional communication through ICT on health outcomes of older adults receiving home care in Japan - a retrospective cohort study. J Gen Fam Med 2022 Jul 20;23(4):233-240 [FREE Full text] [doi: 10.1002/jgf2.534] [Medline: 35800645]
58. Bianchi P, Trimigno M. How does information system success come about in inter-organizational networks of public services? Public Money Manag 2019 Sep 25;41(3):236-245. [doi: 10.1080/09540962.2019.1665361]
59. Bates DW, Samal L. Interoperability: what is it, how can we make it work for clinicians, and how should we measure it in the future? Health Serv Res 2018 Oct;53(5):3270-3277 [FREE Full text] [doi: 10.1111/1475-6773.12852] [Medline: 29527678]
60. Cresswell KM, Worth A, Sheikh A. Integration of a nationally procured electronic health record system into user work practices. BMC Med Inform Decis Mak 2012 Mar 08;12:15 [FREE Full text] [doi: 10.1186/1472-6947-12-15] [Medline: 22400978]
61. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]

Abbreviations

DSCR: digital social care record
ICT: information communication technology
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability
NHS: National Health Service
PRISMA-S: Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by H LaMonica; submitted 03.05.24; peer-reviewed by Y Chu, MP Gagnon; comments to author 07.07.24; revised version received 08.08.24; accepted 13.11.24; published 14.02.25.

Please cite as:
Snow M, Silva-Ribeiro W, Baginsky M, Di Giorgio S, Farrelly N, Larkins C, Poole K, Steils N, Westwood J, Malley J
Best Practices for Implementing Electronic Care Records in Adult Social Care: Rapid Scoping Review
JMIR Aging 2025;8:e60107
URL: https://aging.jmir.org/2025/1/e60107
doi:10.2196/60107
PMID:

Original Paper

Implementation of a Web-Based Program for Advance Care Planning and Evaluation of its Complexity With the Nonadoption, Abandonment, Scale-Up, Spread, And Sustainability (NASSS) Framework: Qualitative Evaluation Study

Doris van der Smissen[1*], PhD; Maud A Schreijer[1*], MSc; Lisette J E W C van Gemert-Pijnen[2], PhD; Rudolf M Verdaasdonk[3], PhD; Agnes van der Heide[1], MD, PhD; Ida J Korfage[1*], PhD; Judith A C Rietjens[1,4*], PhD

[1]Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[2]Psychology, Health & Technology, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, The Netherlands

[3]Health Technology Implementation, Faculty of Science and Technology, University of Twente, Enschede, The Netherlands

[4]Department of Design, Organisation and Strategy, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

[*]these authors contributed equally

Corresponding Author:
Ida J Korfage, PhD
Department of Public Health
Erasmus MC, University Medical Center Rotterdam
Dr. Molewaterplein 40
Rotterdam, 3015 GD
The Netherlands
Phone: 31 107038460
Email: i.korfage@erasmusmc.nl

Abstract

Background: The implementation of eHealth applications often fails. The NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework aims to identify complexities in eHealth applications; the more complex, the more risk of implementation failure.

Objective: This study aimed to analyze the implementation of the web-based advance care planning (ACP) program "Explore Your Preferences for Treatment and Care" using the NASSS framework.

Methods: The NASSS framework enables a systematic approach to improve the implementation of eHealth tools. It is aimed at generating a rich and situated analysis of complexities in multiple domains, based on thematic analysis of existing and newly collected data. It also aims at supporting individuals and organizations to handle these complexities. We used 6 of 7 domains of the NASSS framework (ie, condition, technology, value proposition, adopters, external context, and embedding and adaptation over time) leaving out "organization," and analyzed the multimodal dataset of a web-based ACP program, its development and evaluation, including peer-reviewed publications, notes of stakeholder group meetings, and interviews with stakeholders.

Results: This study showed that the web-based ACP program uses straightforward technology, is embedded in a well-established web-based health platform, and in general appears to generate a positive value for stakeholders. A complexity is the rather broad target population of the program. A potential complexity considers the limited insight into the extent to which health care professionals adopt the program. Awareness of the relevance of the web-based ACP program may still be improved among target populations of ACP and among health care professionals. Furthermore, the program may especially appeal to those who value individual autonomy, self-management, and an explicit and direct communicative approach.

Conclusions: Relatively few complexities were identified considering the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders. Complexities include the broad target population of the program and sociocultural factors. People with limited digital literacy may need support to use the program. Its uptake might be improved by increasing awareness of ACP and the program among a wider population of potential users and among health care professionals. Addressing these issues may guide future use and sustainability of the program.

(JMIR Aging 2025;8:e49507) doi:10.2196/49507

KEYWORDS

eHealth; web-based intervention; implementation; sustainability; advance care planning; NASSS framework; nonadoption, abandonment, scale-up, spread, and sustainability framework; health communication; patient education; patient-centered care

Introduction

eHealth tools can be useful for health promotion, enhancing self-management skills for people with long-term conditions, and patient-physician communication [1]. The web-based program "Explore Your Preferences for Treatment and Care" [2] is aimed at supporting users in engaging in advance care planning (ACP). ACP is a communication process that enables persons to think about their goals and preferences for future treatment and care, to discuss these with their relatives or health care professionals, and to record these if appropriate [3]. The web-based ACP program guides users through 3 steps of ACP—exploration, discussion, and recording of preferences for future treatment and care (Textbox 1). A before and after study showed that the web-based ACP program increased ACP engagement among persons with a chronic disease and was perceived as usable, attractive, and comprehensible [4]. The program was embedded as a decision aid in the Dutch platform "Thuisarts" [5] in April 2020 (English version "GPinfo website" is in development). This platform is owned and hosted by the Dutch College of General Practitioners and provides evidence-based health-related information to the public.

Textbox 1. Main characteristics of the web-based advance care planning program "Explore Your Preferences for Treatment and Care" [2,4].

Content

- Information about advance care planning (ACP).
- Thinking about values and quality of life.
- Communication about preferences with relatives and health care professionals.
- Appointing a health care representative.
- Recording of preferences in an advance directive.
- Reviewing the advance directive.
- References to information about specific diseases, patient organizations, and peer support opportunities.
- The content was based on a scoping review that explored the content, feasibility, and effectiveness of web-based ACP programs [6]; an interview study that identified information needs for web-based ACP of patients with chronic diseases and their relatives [7]; and meetings with a stakeholder group containing relevant stakeholders for the web-based ACP program including patients, relatives, and patient organizations [4].

Structure

- Interactive program; users can watch videos and click on additional information, users are asked questions regarding ACP and can save a document containing their responses.
- Stepwise approach to guide users through the ACP process.
- Embedment in information platform (Thuisarts website [English version: GPinfo website]) which is hosted and owned by the Dutch College of General Practitioners, and provides evidence-based health information.
- Accessible for free.
- Hyperlinks to external websites.
- Text-to-speech option.

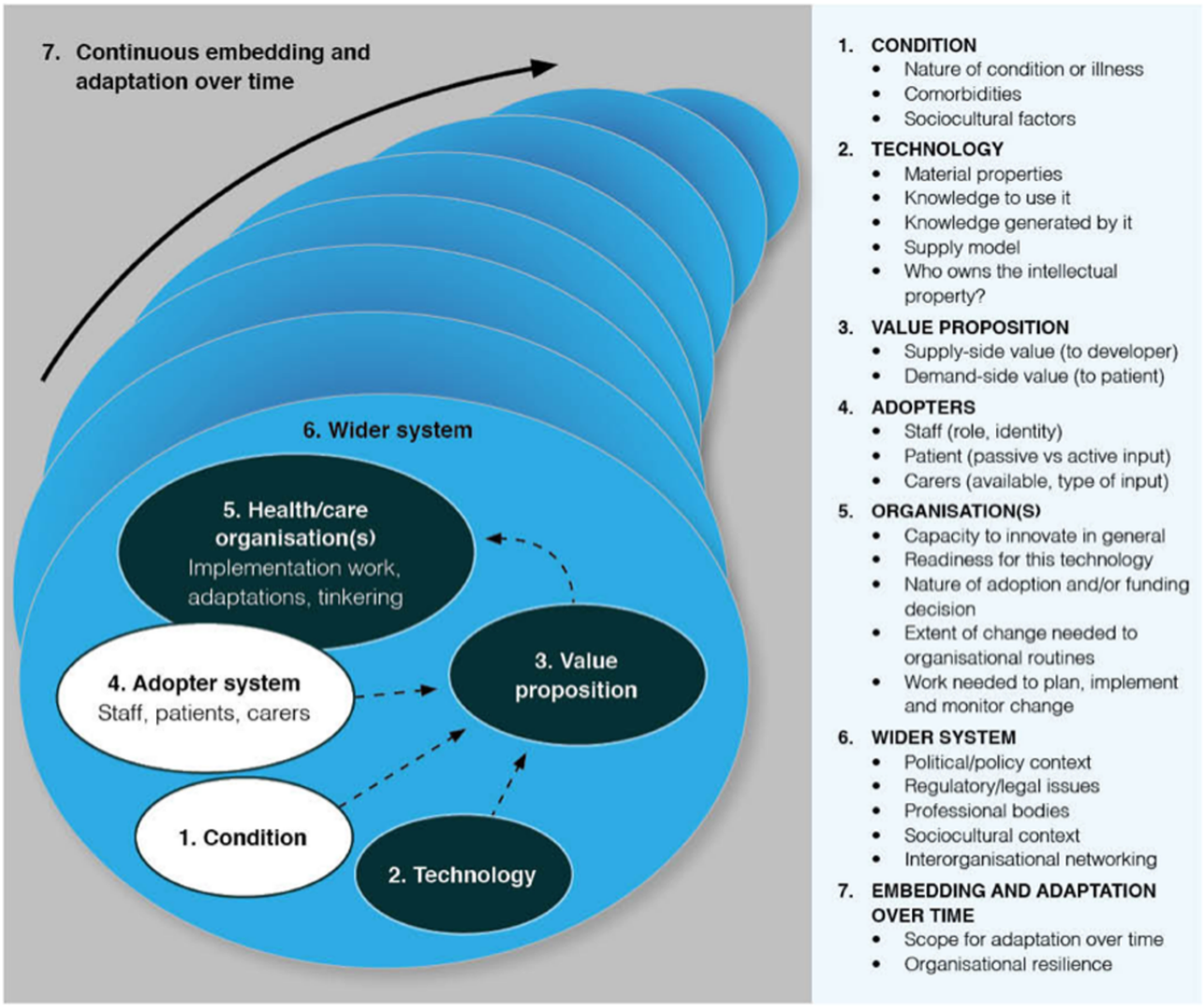
Development

- Developed by researchers (DvdS, IJK, JACR, and AvdH) with expertise in shared decision-making, care at the end of life, and eHealth, in collaboration with stakeholder group including 1 patient, 2 relatives, representatives of the Dutch College of General Practitioners (Nederlands Huisartsen Genootschap), the Dutch Association for Kidney Patients (Nierpatiënten Vereniging Nederland), the Dutch Patient Association (Nederlandse Patiëntenvereniging), Agora (organization to promote the palliative approach), 1 expert in health communication of the Nivel (Netherlands institute for health services research), 1 expert in eHealth of the University of Twente (LvG-P), and 1 representative of Vital Innovators, an organization that conducts Social Return of Investment analyses.
- The development was funded by the Netherlands Health Organisation for Health Research and Development.

Although many eHealth tools are developed, when they are implemented (ie, made available to end users), their sustainable implementation often fails [8]. To identify issues that may hamper the implementation of eHealth tools early, the NASSS (nonadoption, abandonment, scale-up, spread, and sustainability) framework has been developed by Prof T Greenhalgh and her team at Oxford University. The framework is aimed to encourage timely reflecting eHealth applications in health care

and to systematically explore the chances of successful implementation of these eHealth applications by identifying complexities and reducing them [8]. NASSS stands for "nonadoption, abandonment, scale-up, spread, and sustainability over time," which addresses the 5 possible reasons for eHealth implementation to fail [8]. The NASSS framework enables a systematic approach to explore complexities of eHealth tools regarding their implementation; the more complex, the more risk of uptake failure [8]. NASSS focuses at 7 key domains; condition, technology, value proposition, adopters, organization, external context, and embedding and adaptation over time (Figure 1) [8]. A domain can be classified as "simple" when it is straightforward, predictable, and only contains a few components (as in making a sandwich), "complicated" when it contains multiple interacting components (as in building a rocket), and "complex" when it is dynamic, unpredictable, and not easily disaggregated into constituent components (as in raising a child) [8,9]. The NASSS framework postulates that technologies with most domains classified as "simple" are more likely to be successfully implemented than technologies where most domains are classified as "complex" or "complicated." Complexities in different domains can be interdependent, for example, when organizations have difficulty to adopt an eHealth technology, this may complicate adoption of the eHealth tool by staff as well [10,11]. We used the NASSS framework to identify complexities that may hamper sustainable implementation of the web-based ACP program.

Figure 1. The nonadoption, abandonment, scale-up, spread, and sustainability (NASSS) framework [8] published under Creative Commons Attribution 4.0 International License [12].



Methods

Data Analysis

We evaluated the complexity of the web-based ACP program "Explore Your Preferences for Treatment and Care" [2] using the NASSS framework, following the approach of Abimbola et al [11] who also conducted an ex post (retrospective) thematic analysis.

First, we developed a multimodal dataset consisting of different existing, relevant data sources (Table 1). The dataset includes peer-reviewed and published studies describing the development and evaluation of the program [4,6,7], notes of stakeholder group meetings, and a social return on investment analysis in which the social costs (input) and benefits for stakeholders (outcomes) of the program were mapped.

Next, 2 researchers (DvdS an MAS) developed a data extraction form based on the nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool (NASSS-CAT; IRIHS group); the NASSS CAT-LONG interview guide [13]. They used this form to extract relevant pieces of text from the data sources per NASSS domain, that is, (1) condition, (2) technology, (3) value proposition, (4) adopters, (5; previously 6) external context, (6; previously 7) embedding and adaptation over time. We considered the original domain 5, "organization," not applicable because the web-based ACP program is not implemented in a specific health care organization and the hosting organization, the Dutch College of General Practitioners, did not have to change their routines to embed the program. Next, DvdS and MAS together with the wider research group conducted an initial thematic analysis of the extracted information to identify whether sufficient information per domain was available. They conducted 3 additional semistructured interviews to complete the dataset (1) with a representative of the Dutch college of general practitioners (the program host), (2) a general practitioner, and (3) a representative of the organization that funded the development of the web-based ACP program. The interview questions were based on the NASSS CAT-LONG interview guide [13]. Interviewees provided written informed consent for participating in an online interview in which questions were asked about attitudes toward the web-based ACP program related to the domains of the NASSS framework. Interviews were audio-recorded and transcribed verbatim.

Subsequently, DvdS and MAS together with the wider research group conducted a thematic analysis of the final dataset, using the NASSS framework as a lens, leading to a thematic description of complexities per domain. We used relevant ACP literature to further underpin the findings. The conclusion whether a domain is simple, complicated, or complex was based on discussion and eventually consensus among the members of the research group, where we weighed the potential impact and relevance of the different complexities.

Complexities in each of the 6 domains were classified as "simple," "complicated," or "complex." Per domain, all relevant aspects were taken into consideration. An overall classification as "simple" does not require a total absence of room for improvement for a domain; however, most of its elements should be considered as simple. Based on the multimodal dataset and the classification in Multimedia Appendix 1 [8,9], DvdS and MAS prepared a classification per domain, which was discussed during meetings with all researchers (MAS, DvdS, JACR, IJK, and AvdH), until consensus was reached.

Table 1. Overview of the multimodal dataset used to analyze complexities of the web-based advance care planning program "Explore Your Preferences for Treatment and Care."

Sources of information	NASSS[a] framework domains					
	Condition	Technology	Value proposition	Adopters	External context	Embedding and adaptation over time
Results and data from studies:						
Scoping review [6]: • Aim: to determine the feasibility and effectiveness of 11 web-based ACP[b] programs in 27 studies.			✓[c]		✓	
Interview study [7]: • Aim: identifying information needs for web-based ACP. • Participants: 9 patients with a chronic disease and 7 relatives.			✓	✓		
Pilot study [4]: • Aim: assessing the usability and feasibility of the web-based ACP program. • Participants: 3 health care professionals (2 general practitioners and 1 vascular surgeon) and 6 patients with a chronic disease.		✓	✓	✓		
Evaluation study [4]: • Aim: evaluation of ACP engagement (before using the web-based program vs 2 months after completion), usability, and users' satisfaction (including comprehensibility). • Participants: 147 patients with a chronic disease.		✓	✓	✓		
Other available resources:						
Notes of stakeholder group meetings: • Including researchers, patient organizations, patients, and relatives. • Decisions were made concerning the program (eg, definition of target group and content of the program).	✓			✓	✓	✓
SROI[d] analysis: • The social costs and benefits for stakeholders, and mapped investments (input) and revenues (outcomes) of the program. • The SROI methodology followed 9 steps that led to the complete cost-benefit model and the final calculated SROI ratio. • Interviews were conducted with health care organizations, stakeholder group, patients and relatives, and a health insurance company.			✓	✓		✓
Content of the web-based program: • Content, structure, layout, or assembly. • Page with additional information (information about target group and purpose of ACP).	✓	✓				
Platform of the web-based program ("Thuisarts"): • Requirements of the program, consultations with the Thuisarts website. • Guidelines of the Thuisarts website (eg, not saving personal data and maintenance web-based program).	✓	✓		✓	✓	✓

Sources of information	NASSS[a] framework domains					
	Condition	Technology	Value proposition	Adopters	External context	Embedding and adaptation over time
Publicity web-based program: • Interviews with researchers and involved patients and relatives in magazines and the news. • News articles.				✓	✓	✓
• Number of visits to web-based ACP program				✓	✓	
Financial information: • Funder of the development of the web-based ACP program. • Thuisarts website.			✓			
Relevant literature	✓	✓	✓	✓	✓	✓
Newly collected data						
Interviews with 3 additional key stakeholders	✓	✓	✓	✓	✓	✓

[a]NASSS: nonadoption, abandonment, scale-up, spread, and sustainability.

[b]ACP: advance care planning.

[c]✓: included in the dataset.

[d]SROI: social return on investment.

Ethical Considerations

This study was approved by the Medical Research Ethics Committee of the Erasmus MC, University Medical Center Rotterdam on October 21, 2019 (MEC-2019-0590), confirming that the rules laid down in the Medical Research Involving Human Subjects Act do not apply to the interviews as conducted in this study. The study conforms with the International Committee of Medical Journal Editors' recommendations for the conduct, reporting, editing, and publication and for the protection of research participants. The interview participants were provided with information about the aim and content of the interview. They provided written informed consent. Interview participants did not receive compensation. The authors confirm that all patient or personal identifiers have been removed or disguised so the persons described are not identifiable and cannot be identified through the details of the story.

Results

The Condition or Illness

The domain "condition or illness" considers to what extent the condition or illness is well-characterized, well-understood, and predictable and to what extent the condition or illness is influenced by comorbidities or sociocultural factors.

Target Group

The target group of the web-based ACP program is described as follows: "people can engage in ACP when they are healthy, but also when they become older or (chronically) ill, after having an accident or when nearing the end of life" [2]. The target group is hence not strictly demarcated, which is in line with recent recommendations [3,14,15]. During the development of the program, demarcating the target population was a recurring issue with some stakeholders promoting a focus on people with advanced illness and others promoting a broader scope. The target group ambiguity is also visible at the Thuisarts website, where the program is embedded: internal links to the program are included in pages about topics such as end-of-life care and dementia.

Sociocultural Factors Affecting Engagement in ACP

The web-based ACP program is aimed at individuals who are interested in ACP, regardless of their religion or culture. The program contains images and videos of both men (n=5) and women (n=3), persons with various ethnic backgrounds, and a wide range of ages (middle age to older adults) [2]. One video in the program includes a reference to a religious belief.

Nevertheless, the program is currently only available in the Dutch language and was only evaluated among Dutch-speaking persons. Multilingual support is not offered, which limits the accessibility of the program to non–Dutch-speaking individuals. The program fosters autonomy and self-management, and uses an explicit and direct communicative style [2]. This approach may be less appealing to people who consider the role of the family to be central in medical decision-making, and to those who prefer a less direct communication style [16]. This may limit the appeal of the program for those who consider the role of the family to be central in medical decision-making, and to those who prefer a less direct communication style [16].

Classification and Conclusion

The "condition or illness" domain is classified as complex given the rather undemarcated target population of the web-based ACP program and because the program may not address the ACP preferences of all potentially eligible persons.

The Technology

The domain "the technology" encompasses a description of the web-based ACP program and the Thuisarts website, as well as user experiences with the technology.

The Web-Based ACP Program Embedded in Thuisarts Website

The web-based ACP program guides users through the process of ACP in three steps, that are (1) thinking about preferences for future treatment and care, (2) discussing these preferences with relatives and health care professionals, and (3) recording these preferences. Users are asked questions, and the answers to these can be saved in a document and printed. Users of the program are recommended to use this document during a conversation with a relative or a consultation with a health care professional. Users do not need to complete the program at once; they can access it at any preferred moment. The program is freely accessible without the need to register and data of users are not saved. Content maintenance of the program occurs through critical review of the program every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations.

The web-based ACP program is embedded in the Thuisarts website, a web-based platform containing information about health and disease, based on evidence-based guidelines (English version "GPinfo website" is in development). The platform is owned and maintained by the Dutch College of General Practitioners [5]. Approximately two-thirds of the Dutch population are familiar with the platform [17], and it has 6.6 million monthly visitors [18]. The content is written on the B1 level (easily readable and concise), and is offered in standardized formats [19], reading text out loud is enabled, and information is summarized in videos with subtitles. Every year, experts test whether the platform has the recommended level for web-based information (level AA) [20] and identified problems are solved [21].

User Experiences With the Technology

A qualitative pilot study showed that the web-based ACP program was acceptable and feasible for 6 interviewed patients with a chronic disease [4]. In an evaluation study among 147 members of an online research portal (who are expected to have at least some digital skills), including people with levels of low, middle, and high health literacy, the program was perceived as user-friendly (mean score of 70, SD 13, scale 0-100), attractive (mean 3.8, SD 0.7, scale 1-5), and comprehensible (mean 4.2, SD 0.6, scale 1-5) [4]. To support people with limited health- or digital literacy in using the program, people can watch videos and use a text-to-speech option to read the text aloud [4]. Furthermore, a clear and simple structure is used and the number of topics covered is not too large [4]. The texts are written on the B1 level (easily readable and concise) [19].

Classification and Conclusion

The "technology" domain is classified as simple because the web-based ACP program is freestanding and well-maintained, with straightforward, well-understood technology. An evaluation study showed that the program is perceived as user-friendly and comprehensible. However, people with limited health literacy or digital literacy might require help to use the program.

The Value Proposition

The "value proposition" domain concerns whether a new technology is worth developing and for whom it may generate value. It includes demand-side value (value for the users) as well as supply-side value (value for the developers).

Value for Stakeholders

In total, 3 interviewed health care professionals indicated they considered the program as valuable to patients [4] and an evaluation study showed that after use of the web-based ACP program, people contemplated about ACP more often and indicated to feel more ready for ACP [4]. They considered the program to be user-friendly and were satisfied with the program [4]. This is in line with findings from a scoping review, which concluded that participants of web-based ACP programs generally consider these programs as easy to use, not burdensome, and feasible [6]. This review also indicated that web-based ACP programs are effective to improve ACP knowledge, communication about ACP, and documentation of ACP among program users [6].

Regarding the value for health care professionals, the funder of the web-based ACP program project and a general practitioner indicated in the additional interviews to consider the program as helpful to start ACP for health care professionals:

> *Some basic things are in there [in the program] which could be discussed: (…) like what is important to someone? That can help as a guidance in your conversation.* [general practitioner]

Costs

The web-based ACP program was developed in the context of a publicly funded research program and has no commercial purpose. The Dutch College of General Practitioners invests in long-term maintenance of the Thuisarts website and they will maintain the web-based ACP program technically as well.

Cost and Benefit Analysis

The web-based ACP program is freely accessible. The Thuisarts website is not aimed at making a profit. It is financed by the Dutch College of General Practitioners and receives no money from companies seeking financial gain, such as pharmaceutical companies. To better understand the relationship between the total investment (costs) and the expected social effects (benefits) of the web-based ACP program, a so-called social return on investment analysis was conducted [22]. This analysis showed that each investment in terms of money, time, and effort, is expected to generate a social return of approximately 1.7. This means that the program is expected to have a positive value for patients, relatives, health care professionals, and the health care system in terms of money or societal value, such as increased quality of life.

Classification and Conclusion

The "value proposition" domain is classified as simple because the technology is expected to generate a positive value for stakeholders (patients, relatives, health care professionals, and

the health care system) while the benefits of the program are expected to exceed the costs. The Dutch College of General Practitioners invests in long-term maintenance of the Thuisarts website and they will maintain the web-based ACP program technically as well.

The Adopters

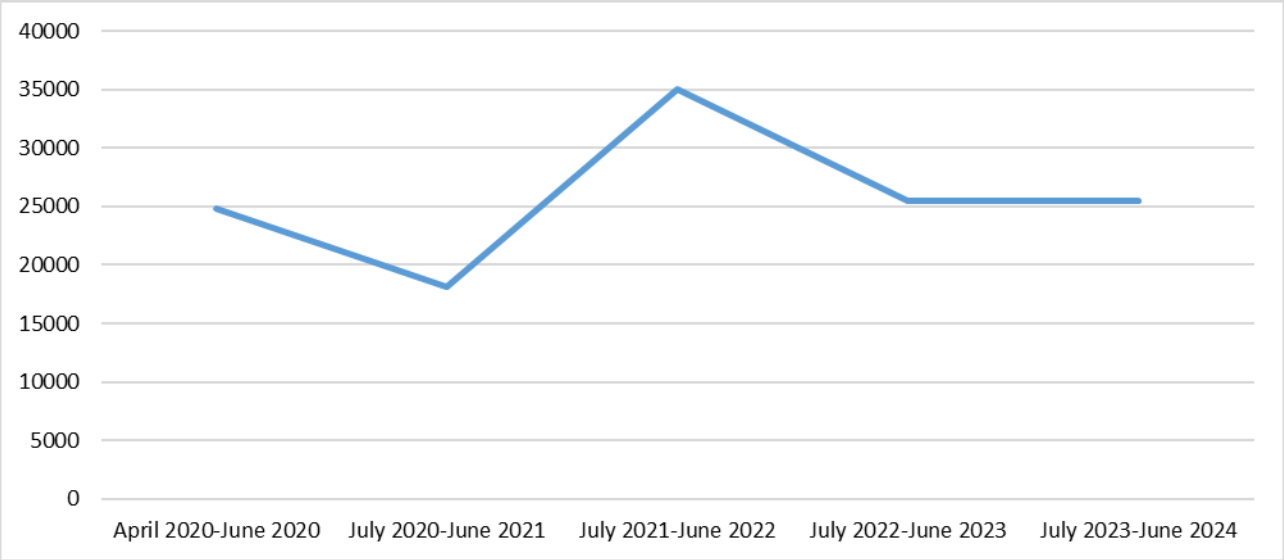
The domain "adopters" focuses on the adoption and continued use of the technology by the groups of people who are intended to actually use the program or to refer potential users to the program, and what they need to use the program or to refer to it. The potential adopters of the web-based ACP program include patients, relatives, health care professionals, and the public in general.

Patients and the Public

Patients with chronic disease and their relatives were asked about their information needs for ACP in a qualitative interview study [7]. They indicated the need for guidance in ACP, information about their disease and care, and information on how to communicate their preferences to relatives and health care professionals [7]. These needs were taken into account in the development of the web-based ACP program. In an evaluation study, users of the web-based ACP program indicated that the program helped them to consider ACP [4]. Users were generally satisfied with the amount of information in the program [4]. A few users mentioned that it was confronting to complete the program because of the difficulty of the topic and a large amount of information. In the first 3 months after its launch (April 1-June 30, 2020) the program was visited 24,849 times. From July 1, 2020, to June 30, 2021, it was visited 18,160 times; from July 1, 2021, to June 30, 2022, it was visited 35,026 times; from July 1, 2022, to June 30, 2023, it was visited 25,423 times; and from July 1, 2023, to June 30, 2024, it was visited 25,510 times. In total, it was visited 128,986 times (Figure 2).

More than 40 medical and patient organizations refer to the web-based ACP program on their websites. However, as described in the domain "condition or illness," not everyone may be interested in engaging in ACP because of a preference to focus on their health and life in the present [23] or valuation of other principles than those central to the web-based ACP program, for example, valuing a central role of the family in medical decision making [16]. While the web-based ACP program is considered useful, it may require substantial commitment and effort of the user in terms of reflecting on their values and preferences for treatment and care.

Figure 2. Visits to the web-based advance care planning program "Explore Your Preferences for Treatment and Care" from its launch (April 1, 2020) to June 30, 2024.



Health Care Professionals

The web-based ACP program recommends participants to involve their health care professionals in the ACP process, for instance, to explain the medical situation, discuss patients' preferences, or file an advance directive. It is known that most people access ACP information through health care professionals [24] and expect that the professional initiates ACP [25]. The Thuisarts website is a frequently used platform among health care professionals; 90% (1751 of 1946) of general practitioners are familiar with it and inform patients about it, and 73% (1421 of 1946) use it for information provision during their consultations [26]. Health care professionals were included in the stakeholder group and provided input during the development of the web-based ACP program. However, opinions about the web-based ACP program of health care professionals have not been extensively evaluated [4]. In total, 3 interviewed health care professionals considered the program to be of value [4]. However, several other studies have described that health care professionals may experience barriers in the initiation of ACP. ACP may be time-consuming, health care professionals report to lack the skills in initiating sensitive ACP conversations [25,27], and they may fear taking away patients' hope [25,27]. Whether the web-based ACP program may support health care professionals to overcome these barriers remains to be studied.

Classification and Conclusion

The "adopters" domain is classified as simple. However, we identified a few complexities. While the web-based ACP

program is potentially useful, it may require substantial commitment and effort of users. Furthermore, it remains to be studied whether the program can support health care professionals to overcome well-known barriers to ACP such as lack of time and fear of taking away patients' hope.

The External Context

The "external context" domain encompasses (1) the policy and political climate, (2) professional organizations and patient organizations, (3) the regulatory context, and (4) the economic context.

Policy and Political Climate

ACP is generally encouraged by organizations of health care professionals in the Netherlands, such as general practitioners, home care organizations, and hospitals. Many initiatives are taken to implement ACP in health care practice such as the development of ACP tools and manuals, websites for the public as well as health care professionals such as Palliaweb website, and ACP modules for electronic patient files [3,28-30]. In addition, the Dutch Ministry of Health, Welfare and Sport has spent €51 million (US $52.6 million) in the period 2014-2020 on research on palliative care including ACP [30] and initiated another research program on ACP in 2022 [31]. Furthermore, the Ministry has organized a public information campaign in 2022 to raise awareness for the importance of timely talking about death and dying [32,33].

Professional Organizations and Patient Organizations

Several patient organizations were involved in the development, evaluation, and implementation of the web-based ACP program. More than 40 organizations, including The Netherlands Patient Federation and the Dutch Association for Kidney Patients, as well as general practitioners and care organizations, referred to the web-based ACP program on their websites and in newsletters.

Regulatory Context

The regulatory context of ACP in general in the Netherlands is supportive regarding ACP. Several professional bodies have developed and provided guidelines for ACP [34], toolkits [35], and manuals for doctors [29,36]. Some of these specifically refer to the web-based ACP program [29,36]. Some materials are specifically developed for patients [37]. However, these ACP materials have different target groups, for instance, health care providers [29,34,36], patients in the palliative phase or at the end of life [3,14,15], or broader populations [3], and therefore it may not always be clear to whom they apply. Regarding the regulatory context of the web-based ACP program, users' privacy is protected in compliance with data security and privacy requirements, by not requiring any log-in procedure and not recording any details of users.

Economic Context

The reimbursement of ACP for health care professionals is not clearly regulated.

Classification and Conclusion

The "external context" domain is classified as simple. Professional bodies are supportive toward the web-based ACP program and ACP in general. We found some complexities, such as different available manuals and tools for ACP containing different recommendations, and the unavailability of a treatment code for health care professionals. However, these external conditions are not likely to complicate the adoption of the web-based ACP program to a large extent.

Embedding and Adaptation Over Time

Overview

The scope of the Thuisarts website is expanding. Information provision was focused on primary care in the past, but currently also includes information on secondary care. The Thuisarts website, including the web-based ACP program, is likely to be maintained in the future, with both technical and content updates. Content maintenance of the program occurs through a critical review of the program every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations. They will review compliance with (evolving) data security regulations, privacy requirements, and policies. If necessary, the program will be adapted. At this time, no changes in data protection regulations or policies are expected in the near future that would affect the program.

Our evaluation study showed that in general, participants had a positive user experience and were satisfied with the web-based ACP program [4]. To further improve this, we are systematically collecting informal feedback on the program and visitor numbers. These will be discussed during each 3-yearly review of the program by the stakeholder group. If necessary, they will make recommendations to adapt the program. To do so, the web-based ACP program will be adapted using an iterative design strategy of ongoing collecting informal feedback and the editorial team of the Thuisarts website adapting the program accordingly. Based on the feedback, user experience and satisfaction are aimed to be enhanced by improving the usability and feasibility of the web-based program.

Classification and Conclusion

We classify the "embedding and adaptation over time" domain as simple, since it is unlikely that the web-based ACP program as embedded in the Thuisarts website or its value are significantly going to change in the next years.

Discussion

Principal Results

To analyze the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care," the NASSS framework was used to identify complexities within 6 domains. The analysis revealed no or little complexity in the domains "technology," "value proposition," "external context," and "embedding and adaptation over time." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders as well as for the Thuisarts website. The analyses revealed some complexities in the "condition" domain, including the broad and rather undemarcated target population of ACP and the web-based ACP program and sociocultural factors that may

limit users' engagement. In the "adopters" domain, complexity entails that the program requires substantial commitment and effort by users with respect to reflecting on their preferences for future treatment and care. Some people with limited digital literacy may need support to use the program. Furthermore, it is yet unstudied to what extent health care professionals adopt the web-based ACP program. The program is embedded in the general practitioners' platform "Thuisarts," which is frequently used by general practitioners and the general public. The fact that the web-based ACP program is not embedded in a specific health care organization simplified its implementation, since no alignment regarding content or layout was required. On the other hand, embedding the program in the routine practice of a health care organization, could have boosted its use. Finally, we found some complexity in the "external context" domain. While the political and policy climate, regulatory context, professional organizations, and patient organizations are generally supportive regarding ACP, a national guideline for ACP and a treatment code to reimburse ACP could further increase willingness and ability of health care providers to engage in ACP. Overall, the analysis showed that the program has good potential for sustainable implementation, as it is expected to be continuously used in practice [8,38,39], will be updated if necessary, and will be maintained long-term.

Comparison With Previous Work

The results of this study showed that the web-based ACP program "Explore Your Preferences for Treatment and Care" has relatively few complexities regarding the implementation. This may have several reasons. First, the user-centered design of the program included an extensive preparatory phase, including a scoping review to learn from existing web-based ACP programs [6], identification of the needs of patients with chronic disease and their relatives for web-based ACP in an interview study [7], and close collaboration with stakeholders including patients and relatives. Second, during the development of the program, we paid due attention to sustainable implementation in an existing, well-used platform—"Thuisarts." The team of the Thuisarts website has expertise in developing information for the public with different levels of health literacy. It is remarkable that other web-based ACP programs often seem to have not paid due attention to implementation. Three recent reviews about web-based ACP programs [6,40,41] showed that studies have mainly focused on the development of web-based ACP programs and evaluation of their feasibility, usability, acceptability, and effectiveness, but most did not address their implementation. This is in line with reviews about implementing eHealth in general showing that research has mainly focused on the content and evaluation of eHealth tools, but less so took into account the external context and sustainable adoption in health care systems [8,10].

The results of the current study also indicated that health care professionals are an important group of adopters for the web-based ACP program. The web-based ACP program, following existing guidelines, recommends an active role of health care professionals in the ACP process of patients, for instance, to discuss diagnosis, prognosis, and relevant treatment and care options with the patient, and to document the ACP process in the medical file [3,4]. In the literature, several barriers are described that health care professionals may experience in the initiation of ACP, such as lack of time, lack of skills in conducting ACP conversations, and fear of taking away patients' hope [25,27]. In the development of the web-based ACP program, the role of health care professionals was to a limited extent taken into account. First, the program is embedded in the general practitioners' platform "Thuisarts," which is frequently used by general practitioners [26] and the general public [17,18]. Furthermore, health care professionals were included in the stakeholder group and provided input during the development of the web-based ACP program. In the program, the user is encouraged to discuss and share their preferences with their health care professionals at several points [2]. However, the actual use of the web-based ACP program among health care professionals and their views and experiences regarding the program, are unstudied. Given the barriers that health care professionals may experience in the initiation of ACP [25,27], a patient-centered and community approach of ACP is important, in which patients can also take the initiative to engage in ACP themselves. This is central to the program, as the program supports the initiation of ACP, self-management of users, and encourages conversation with health care professionals about values, goals, and preferences.

The results indicated complexities considering the broad target population of ACP in the web-based ACP program. This complexity is a reflection of the evolving concept of ACP in the literature. Initially, ACP was conceptualized to be used for patients in the palliative phase and at the end of life [3,14,15]. More recently, an international consensus study on ACP recommended that individuals can engage in ACP in any stage of life, but that ACP can be more targeted when individuals' health condition worsens or when their age increases [3]. This evolving concept of ACP aligns with the evolving concept of palliative care, whose definition has widened from a sole focus on the end of life to also including chronic illnesses [42] and a shift from in-hospital to community-based care [43]. Reaching the community with high-quality information and guidance about ACP through the Thuisarts website may therefore fit these developments.

Strengths and Limitations

A strength of this study is our systematic approach of evaluating the web-based ACP program using the NASSS framework. It provided insight in its complexities that might hamper its implementation. Furthermore, we had a large dataset including publications about the development and evaluation of the web-based ACP program, notes of stakeholder meetings, and interviews with the program host, a general practitioner, and a funding agency representative. Insight into the complexities may be used to improve future sustainable implementation of the program, and may be used to improve other web-based programs as well.

A limitation is the start of the NASSS evaluation after the tool was developed and embedded in the Thuisarts website, which has limited the opportunity to adapt the program during its development. Some of the authors (DvdS, JACR, AvdH, and IJK) were the developers of the web-based ACP program, and therefore they could provide in-depth insight in the

developmental process. However, although the authors aimed to analyze the implementation of the web-based ACP program objectively, there can be a bias in the evaluation.

Recommendations to Improve Complexities of the Web-Based Advance Care Planning Program

To enhance sustainable implementation, we provide several recommendations to improve the identified complexities of the web-based ACP program.

First, to make the program available in other languages than Dutch. This may further enhance the accessibility of the program to people with different languages. For instance, an English version of the program could be placed on the GPinfo website, which is the English version of the Thuisarts website that is still in development. The web-based ACP program may be translated to other languages as well, to enhance its accessibility to people who are not proficient in Dutch. To improve cultural inclusiveness, future research should explore the needs of people from various cultural backgrounds with regard to ACP and if desired, the program could be culturally adapted or different versions of the program could be developed.

Second, to enhance accessibility of the web-based ACP program to persons with very limited health or digital literacy. The program was developed to assist individuals with limited health or digital literacy, by offering readable texts, text-to-speech options, and videos. However, individuals with very limited health or digital literacy may still experience barriers to use the program. To enhance accessibility of the program for these users, recommendations to complete the program together with a relative could be added to the program. Furthermore, a PDF version of the web-based ACP program could be developed, that could be printed by a relative or health care professional, to make the program also available "offline."

Third, to inform health care professionals about the variance in readiness for ACP among patients. Not everyone may want to engage in ACP [44], for instance, due to different cultural backgrounds and different needs to consider treatment and care preferences. In addition, we recommend to inform and support health care professionals in how to conduct ACP with their patients. The web-based ACP program may help to overcome lack of time, since health care professionals can refer to the program, patients can prepare themselves for an ACP discussion on beforehand, and then the discussion with the health care professional can be scheduled.

Fourth, the concept of ACP as an ongoing process of considering preferences (also including healthy individuals) [3], could be embedded in national guidelines, and be routinely integrated in health care practice to reach a broader target group for ACP. National guidelines and availability of a treatment code may lower barriers in ACP for health care professionals, and may consequently lead to increased engagement with ACP for patients. Websites aimed at the general public could refer to the web-based ACP program, for instance, pages about shared decision-making and pages within the Thuisarts website about chronic diseases. Furthermore, we recommend to conduct public awareness campaigns about ACP (including web-based ACP programs), as these have been found to enhance awareness of ACP [45]. These may also help to enhance accessibility of ACP for people beyond the medical health care setting.

Finally, to ensure adaptation of the program to evolving data protection requirements. During critical review of the program, that will take place every 3 years by a stakeholder group of patients, relatives, health care professionals, and patient organizations, compliance with (evolving) data security regulations, privacy requirements, and policies should be carefully reviewed. If necessary, the program should be adapted. In addition, the strategy of the Thuisarts website, the host of the web-based ACP program, is aimed at adapting the information on the Thuisarts website to the evidence-based guidelines for general practitioners as developed by the Dutch College of General Practitioners. When guidelines change, the information at the Thuisarts website is adapted accordingly by their editorial team. Furthermore, the strategy of the Thuisarts website focuses on collaborating with health care professionals and policy makers, and on including patients' perceptions in the information, to ensure the content is adapted to significant changes in health care landscapes or policies.

Since the domains of the NASSS framework may be interdependent, reducing complexity in a particular domain may consequently also reduce the complexity in other domains. For instance, in case of national guidelines on ACP, health care professionals may be more willing to initiate ACP with their patients and patients may be more ready to initiate ACP.

Recommendations for Future Research

The following recommendations, based on the results and conclusions of the current study, can aid future implementation of comparable web-based ACP programs.

In order to improve the sustainable implementation of eHealth tools, in particular web-based ACP programs, first, we recommend to systematically and throughout the process (from idea to implementation) assess their complexities, with careful attention for all potential adopters, their external context, and their embedment and adaptation over time. The NASSS framework is suitable to conduct such evaluation, as well as other models that can be used for implementation or sustainability evaluation, such as the Normalization Process Theory (a framework to understand how interventions become embedded or "normalized" in health care settings) [46] and the Centre for eHealth Research Roadmap (CeHRes) [47,48]. These frameworks consider similar domains as the NASSS framework, sometimes ordered slightly differently, and can provide guidance for the sustainable development and implementation of eHealth technologies [46-48]. We expect that others who aim to implement similar web-based ACP programs and who apply comparable implementation frameworks, will benefit from the findings of this study and the recommendations. Yet, these should be contextualized for each application (eg, aligned with local policy frameworks and targeted to the target groups central in that program).

Second, we recommend identifying all relevant target groups of the eHealth tools, and explore their needs and preferences to ensure the tool meets their needs. For the web-based ACP program in particular, we recommend exploring health care

professionals' perspectives as well as their actual use and awareness of the program.

Third, we recommend assessing as well as improving the attractiveness of eHealth tools to potential users, including a culturally diverse population and to groups with low digital literacy.

Fourth, we recommend that research includes a long-term perspective regarding the implementation of eHealth tools. This may include evaluating the long-term effects as well as the identification of new scientific evidence that necessitates an update of the content of eHealth tools.

Conclusions

Relatively few complexities were identified considering the implementation of the web-based ACP program "Explore Your Preferences for Treatment and Care." The program is evidence-based, freestanding, and well-maintained, with straightforward, well-understood technology. The program is expected to generate a positive value for different stakeholders. Complexities include the broad target population of the program and sociocultural factors. People with limited digital literacy may need support to use the program. Its uptake might be improved by increasing awareness of ACP and the program among a wider population of potential users and among health care professionals. Addressing these issues may guide future use and sustainability of the program.

Acknowledgments

The authors would like to thank the interviewees for participating in this study. The authors would also like to thank the developers of the NASSS framework Prof Dr Trish Greenhalgh and Dr Chrysanthi Papoutsi for their valuable insights regarding the NASSS framework and their permission to reproduce the original NASSS Figure 1 and Multimedia Appendix 1.

This work was supported by the Netherlands Organisation for Health Research and Development (grant 555003026). The funding source had no involvement in the study design, collection, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication. The authors declare that no generative artificial intelligence was used in any portion of the manuscript writing.

Data Availability

Research data will be available for verification for at least 15 years. All data will be retained and stored at the central network server of the Erasmus MC, University Medical Center Rotterdam, and can be accessed by those specifically granted access.

Authors' Contributions

All authors contributed to the concept or design of the work or acquisition, analysis, or interpretation of data; drafted the paper or revised it critically for important intellectual content; approved the version to be published; and have participated sufficiently in the work to take public responsibility for appropriate portions of the content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Domains and questions in the Non-adoption, abandonment, scale-up, spread, and sustainability (NASSS) framework (reproduced from Greenhalgh T et al [8], published under Creative Commons Attribution 4.0 International License [12].
[PDF File (Adobe PDF File), 165 KB - aging_v8i1e49507_app1.pdf]

References

1. Murray E. eHealth: where next? Br J Gen Pract 2014;64(624):325-326 [FREE Full text] [doi: 10.3399/bjgp14X680365] [Medline: 24982470]
2. Thuisarts.nl. Verken uw wensen voor zorg en behandeling. The Dutch College of General Practitioners. 2020 Apr 01. URL: https://www.thuisarts.nl/keuzehulp/verken-uw-wensen-voor-zorg-en-behandeling [accessed 2022-10-11]
3. Rietjens JAC, Sudore RL, Connolly M, van Delden JJ, Drickamer MA, Droger M, European Association for Palliative Care. Definition and recommendations for advance care planning: an international consensus supported by the European association for palliative care. Lancet Oncol 2017;18(9):e543-e551 [FREE Full text] [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
4. van der Smissen D, Rietjens JAC, van Dulmen S, Drenthen T, Vrijaldenhoven-Haitsma FRMD, Wulp M, et al. The web-based advance care planning program "Explore Your Preferences for Treatment and Care": development, pilot study, and before-and-after evaluation. J Med Internet Res 2022;24(12):e38561 [FREE Full text] [doi: 10.2196/38561] [Medline: 36459410]

5. Thuisarts.nl. The Dutch College of General Practitioners. Over Thuisarts.nl. 2022. URL: https://www.thuisarts.nl/over-thuisartsnl [accessed 2022-10-11]
6. van der Smissen D, Overbeek A, van Dulmen S, van Gemert-Pijnen L, van der Heide A, Rietjens JA, et al. The feasibility and effectiveness of web-based advance care planning programs: scoping review. J Med Internet Res 2020;22(3):e15578 [FREE Full text] [doi: 10.2196/15578] [Medline: 32181750]
7. van der Smissen D, Rietjens JAC, van Gemert-Pijnen L, van Dulmen S, van der Heide A, Korfage IJ. Information needs of patients with chronic diseases and their relatives for web-based advance care planning: a qualitative interview study. BMC Palliat Care 2021;20(1):77 [FREE Full text] [doi: 10.1186/s12904-021-00770-x] [Medline: 34053435]
8. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
9. Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. BMJ 2001;323(7313):625-628 [FREE Full text] [doi: 10.1136/bmj.323.7313.625] [Medline: 11557716]
10. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629 [FREE Full text] [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
11. Abimbola S, Patel B, Peiris D, Patel A, Harris M, Usherwood T, et al. The NASSS framework for ex post theorisation of technology-supported change in healthcare: worked example of the TORPEDO programme. BMC Med 2019;17(1):233 [FREE Full text] [doi: 10.1186/s12916-019-1463-x] [Medline: 31888718]
12. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2025-01-30]
13. Greenhalgh T, Maylor H, Shaw S, Wherton J, Papoutsi C, Betton V, et al. The NASSS-CAT tools for understanding, guiding, monitoring, and researching technology implementation projects in health and social care: protocol for an evaluation study in real-world settings. JMIR Res Protoc 2020;9(5):e16861 [FREE Full text] [doi: 10.2196/16861] [Medline: 32401224]
14. Ott B, van Thiel GJ, de Ruiter CM, van Delden HJ. [Timing of advance care planning in frail elderly patients: when to start?] Kwetsbare ouderen en advance care planning: Wanneer beginnen? Ned Tijdschr Geneeskd 2015;159:A8295. [Medline: 25650032]
15. Billings JA, Bernacki R. Strategic targeting of advance care planning interventions: the Goldilocks phenomenon. JAMA Intern Med 2014;174(4):620-624. [doi: 10.1001/jamainternmed.2013.14384] [Medline: 24493203]
16. Cheng SY, Lin CP, Chan HYL, Martina D, Mori M, Kim SH, et al. Advance care planning in Asian culture. Jpn J Clin Oncol 2020;50(9):976-989. [doi: 10.1093/jjco/hyaa131] [Medline: 32761078]
17. Bijna tweederde Nederlanders bekend met Thuisarts.nl. ICTHealth.nl. 2021. URL: https://www.icthealth.nl/nieuws/bijna-twee-derde-nederlanders-bekend-met-thuisarts-nl/ [accessed 2022-07-21]
18. Forse stijging aantal bezoeken Thuisarts.nl. Federatie Medisch Specialisten. 2022. URL: https://demedischspecialist.nl/nieuwsoverzicht/nieuws/forse-stijging-aantal-bezoeken-thuisartsnl [accessed 2023-02-15]
19. Springvloet L, Bos N, Jong J, Friele R, Boer D. De Transparantiemonitor 2018. Thuisarts.nl: Hoe draagt Thuisarts.nl bij aan transparantie en de best passende zorg?. 2019. URL: https://www.nivel.nl/nl/publicatie/de-transparantiemonitor-2018-thuisartsnl-hoe-draagt-thuisartsnl-bij-aan-transparantie-en [accessed 2025-01-03]
20. DigiToegankelijk. Logius, ministerie van binnenlandse zaken en koninkrijkrelaties. WCAG: Richtlijnen voor webcontent. 2022. URL: https://www.digitoegankelijk.nl/uitleg-van-eisen/wat-wcag [accessed 2022-12-02]
21. Thuisarts. Thuisarts.nl ook bruikbaar voor mensen met beperking. The Dutch College of General Practitioners. URL: https://www.thuisarts.nl/thuisartsnl-ook-bruikbaar-voor-mensen-met-beperking [accessed 2022-10-11]
22. Nicholls J, Lawlor E, Neitzert E, Goodspeed T. A Guide to Social Return on Investment. London: The Cabinet Office; 2012.
23. McIlfatrick S, Slater P, Bamidele O, Muldrew D, Beck E, Hasson F. 'It's almost superstition: If I don't think about it, it won't happen'. Public knowledge and attitudes towards advance care planning: A sequential mixed methods study. Palliat Med 2021;35(7):1356-1365 [FREE Full text] [doi: 10.1177/02692163211015838] [Medline: 34000901]
24. Rodi H, Detering K, Sellars M, Macleod A, Todd J, Fullerton S, et al. Exploring advance care planning awareness, experiences, and preferences of people with cancer and support people: an Australian online cross-sectional study. Support Care Cancer 2021;29(7):3677-3688. [doi: 10.1007/s00520-020-05878-z] [Medline: 33184712]
25. Jabbarian LJ, Zwakman M, van der Heide A, Kars MC, Janssen DJA, van Delden JJ, et al. Advance care planning for patients with chronic respiratory diseases: a systematic review of preferences and practices. Thorax 2018;73(3):222-230. [doi: 10.1136/thoraxjnl-2016-209806] [Medline: 29109233]
26. Drenthen T, Beijaert RPH, Jansen PWM, Korevaar JC, Smeele IJM. [What do you think of Thuisarts.nl? Experiences after 3 years of www.Thuisarts.nl] Thuisarts.nl, hoe bevalt dat? Ervaringen na 3 jaar Thuisarts.nl. Ned Tijdschr Geneeskd 2014;158:A8282. [Medline: 25515391]
27. Jimenez G, Tan WS, Virk AK, Low CK, Car J, Ho AHY. Overview of systematic reviews of advance care planning: Summary of evidence and global lessons. J Pain Symptom Manage 2018;56(3):436-459.e25 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.05.016] [Medline: 29807158]

28. Palliaweb. Advance care planning. Palliaweb. 2022. URL: https://palliaweb.nl/zoeken?searchText=advance+care+planning [accessed 2022-12-12]
29. Boddaert M, Caffarel J, Dermois M, Kodde A, Reyners AKL. Leidraad voor het proces en uniform vastleggen van proactieve zorgplanning (advance care planning, acp) naar aanleiding van de covid-19-pandemie. 2020. URL: https://palliaweb.nl/getmedia/1e1ce663-6f7d-46ac-a47a-c8cb589597b7/Leidraad-proactieve-zorgplanning-(ACP)_22102020_new.pdf [accessed 2020-10-22]
30. Programma palliantie. Meer dan zorg. ZonMw. 2022. URL: https://www.zonmw.nl/nl/over-zonmw/onderwijs/programmas/programma-detail/palliantie-meer-dan-zorg/ [accessed 2022-12-12]
31. ZonMw. Inzet van proactieve zorgplanning in de praktijk (uitgewerkte subsidieaanvragen). ZonMw. 2022. URL: https://www.zonmw.nl/nl/subsidies/openstaande-subsidieoproepen/detail/item/inzet-van-proactieve-zorgplanning-in-de-praktijk-uitgewerkte-subsidieaanvragen/ [accessed 2023-03-03]
32. SIRE. Campagne 'de dood, praat erover' campaign 'death, talk about it'. URL: https://sire.nl/campagnes/de-dood-praat-erover/ [accessed 2022-10-27]
33. Rijksoverheid. Rijksoverheid [Dutch Government]. Campagnetoolkits palliatieve zorg. 2019. URL: https://www.campagnetoolkits.nl/palliatieve-zorg [accessed 2022-10-27]
34. Rietjens J, Sudore R, Deliens L, Ott B, Connolly M, van Delden J, et al. Internationale definitie van advance care planning. Huisarts Wet 2018;61(8):22-29 [FREE Full text] [doi: 10.1007/s12445-018-0224-y]
35. Toolkit advance care planning mbt het levenseinde. Frank Guldemond, Brenda Ott, Annet Wind, kaderhuisartsen ouderengeneeskunde. 2017. URL: https://www.laego.nl/wp-content/uploads/2018/04/Toolkit-Advance-Care-Planning-2017.pdf [accessed 2022-12-12]
36. KNMG. Patiëntenfederatie Nederland. Handreiking voor artsen om met patiënten in gesprek te gaan tijdig praten over het levenseinde. KNMG and Patiëntenfederatie Nederland. URL: https://data.maglr.com/3230/issues/30306/393465/downloads/handreiking-tijdig-praten-over-het-levenseinde.pdf [accessed 2022-12-12]
37. KNMG. Patiëntenfederatie Nederland. Praat op tijd over je levenseinde over passende zorg in je laatste levensfase KNMG and Patiëntenfederatie Nederland. URL: https://www.knmg.nl/advies-richtlijnen/dossiers/praten-over-het-levenseinde-1.htm [accessed 2022-12-12]
38. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013;8:117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
39. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011;38(1):4-23 [FREE Full text] [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]
40. Dupont C, Smets T, Monnet F, Pivodic L, De Vleminck A, Van Audenhove C, et al. Publicly available, interactive web-based tools to support advance care planning: Systematic review. J Med Internet Res 2022;24(4):e33320 [FREE Full text] [doi: 10.2196/33320] [Medline: 35442207]
41. Butler M, Ratner E, McCreedy E, Shippee N, Kane RL. Decision aids for advance care planning: an overview of the state of the science. Ann Intern Med 2014;161(6):408-418 [FREE Full text] [doi: 10.7326/M14-0644] [Medline: 25069709]
42. Radbruch L, De Lima L, Knaul F, Wenk R, Ali Z, Bhatnaghar S, et al. Redefining palliative care-A new consensus-based definition. J Pain Symptom Manage 2020;60(4):754-764 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.04.027] [Medline: 32387576]
43. Sibbald B, McDonald R, Roland M. Shifting care from hospitals to the community: A review of the evidence on quality and efficiency. J Health Serv Res Policy 2007;12(2):110-117. [doi: 10.1258/135581907780279611] [Medline: 17407662]
44. Zwakman M, Jabbarian LJ, van Delden J, van der Heide A, Korfage IJ, Pollock K, et al. Advance care planning: a systematic review about experiences of patients with a life-threatening or life-limiting illness. Palliat Med 2018;32(8):1305-1321 [FREE Full text] [doi: 10.1177/0269216318784474] [Medline: 29956558]
45. Seymour J. The impact of public health awareness campaigns on the awareness and quality of palliative care. J Palliat Med 2018;21(S1):S30-S36 [FREE Full text] [doi: 10.1089/jpm.2017.0391] [Medline: 29283867]
46. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization process theory. Implement Sci 2009;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
47. van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R. Ehealth research, theory and development: A multi-disciplinary approach. Oxfordshire: Routledge; 1999.
48. CEHRES Roadmap Toolkit. Centre for eHealth & Wellbeing Research, University of Twente. 2018. URL: https://www.utwente.nl/en/bms/ehealth/cehres-roadmap-toolkit/ [accessed 2023-05-18]

Abbreviations

ACP: advance care planning
CeHRes: Centre for eHealth Research Roadmap
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability

NASSS-CAT: nonadoption, abandonment, scale-up, spread, and sustainability complexity assessment tool

Edited by Y Jiang; submitted 01.06.23; peer-reviewed by F Monnet, P Delir Haghighi, Y Chu; comments to author 25.10.23; revised version received 16.02.24; accepted 06.11.24; published 04.03.25.

Please cite as:
van der Smissen D, Schreijer MA, van Gemert-Pijnen LJEWC, Verdaasdonk RM, van der Heide A, Korfage IJ, Rietjens JAC Implementation of a Web-Based Program for Advance Care Planning and Evaluation of its Complexity With the Nonadoption, Abandonment, Scale-Up, Spread, And Sustainability (NASSS) Framework: Qualitative Evaluation Study
JMIR Aging 2025;8:e49507
URL: https://aging.jmir.org/2025/1/e49507
doi:10.2196/49507
PMID:40053753

Original Paper

Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study

Siow-Hooi Tan[1], PhD; Yee-Yann Yap[1], MPhil; Siow-Kian Tan[2], PhD; Chee-Kuan Wong[3], MD, MRCPE

[1]Faculty of Management, Multimedia University, Cyberjaya, Malaysia

[2]School of Economics and Management, Xiamen University Malaysia, Sepang, Malaysia

[3]Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:
Siow-Hooi Tan, PhD
Faculty of Management, Multimedia University
Persiaran Multimedia
Cyberjaya, 63100
Malaysia
Phone: 60 38312 ext 5658
Email: shtan@mmu.edu.my

Abstract

Background: The aging population and the accompanying rise in chronic diseases have intensified the need to study the adoption of telehealth services. However, the success of telehealth services depends not only on their ease and usefulness but also on addressing broader concerns. Despite being a substantial user group in traditional health services, older adults may encounter barriers to adopting telehealth services. Increasing the adoption of telehealth among the older adult population is crucial for enhancing their access to care and managing the challenges of aging effectively.

Objective: We aimed to explore factors influencing the adoption of telehealth services among older adults in Malaysia, going beyond the conventional framework by incorporating transition cost and subjective well-being as additional constructs.

Methods: A cross-sectional survey was conducted among 119 adults aged ≥60 years in Malaysia, using 39 survey items adapted from existing studies. Data analysis was performed using partial least squares structural equation modeling, with both the measurement model and structural model being evaluated. To determine the predictive relevance of the model, PLSpredict was applied. In addition, importance-performance map analysis was conducted to further expand on the structural model results by assessing the performance of each variable.

Results: Of the 119 participants, 52 (43.7%) were women and 67 (56.3%) were men. The study found that subjective well-being (β=0.448; *P*<.001) was the most significant factor, followed by attitude (β=0.242; *P*<.001), transition cost (β=−0.163; *P*<.001), and perceived usefulness (β=0.100, *P*=.02) in influencing telehealth service intention. Furthermore, perceived ease of use (β=0.271; *P*<.001), availability (β=0.323; *P*<.001), subjective well-being (β=0.261; *P*<.001), and trust (β=0.156, *P*=.004) positively influenced perceived usefulness, while inertia (β=0.024, *P*=.22) did not. In addition, availability (β=0.420; *P*<.001) and subjective well-being (β=0.260; *P*<.001) were positively related to perceived ease of use, with inertia (β=−0.246; *P*<.001) having a negative impact. The importance-performance map analysis results showed that subjective well-being (importance=0.532) was the most crucial factor for older adult users, while availability (importance=70.735) had the highest performance in telehealth services.

Conclusions: This research underscores the importance of catering to the subjective well-being of older adults and optimizing the availability of telehealth services to encourage adoption, ultimately advancing health care accessibility and quality for this vulnerable demographic.

(JMIR Aging 2025;8:e60936) doi:10.2196/60936

KEYWORDS

telehealth services adoption; survey; questionnaire; telehealth; older adult population; subjective well-being; transition cost; technology acceptance model; importance-performance map analysis; IPMA

Introduction

Background

Telehealth refers to the delivery of long-distance clinical health care services by health care professionals using electronic information and telecommunications technologies. Due to the growth of the internet and communication infrastructure, telehealth has gradually developed into a practical and secure way for patients to obtain reliable information and medical consultation [1]. Using telehealth has several advantages, such as eliminating the need for direct patient–health care provider interaction during regular treatment. Telehealth can also provide remote care, which can reduce the need for medical center resources and increase the accessibility of care.

Previous telehealth studies have developed various concepts to address how telehealth could fulfill the needs of older adults, such as in the context of chronic disease management, enhancing independent living and improving their overall well-being [2]. Telehealth can be useful for older adults with chronic diseases to monitor their conditions at home. For instance, with the use of telehealth, older adult patients can prevent unnecessary hospitalization and still ensure they receive emergency treatment in a cost-efficient manner [3,4]. In addition, Chou et al [5] found that telehealth can improve the well-being of older adults by enhancing their quality of life. Furthermore, telehealth can promote independent living at home among older adults [6].

Older adults could benefit from telehealth as they are the fragile groups who may need this service sooner or later. Nevertheless, they are also the group most concerned about technology. The literature provides evidence that older adults can receive several advantages with the use of telehealth services, such as health monitoring and care, disease prevention, improved quality of life, and independent living. Despite all the advantages, the adoption of telehealth technology may be challenging for older adults because they are slower and more resistant to adopting new technology as they tend to be more traditional, cautious, risk-averse, and suspicious toward innovations [7]. There is a lack of knowledge about what factors individuals will consider when accepting telehealth [8]. The market response indicates that acceptance of technology by older adults is a complex problem that is impacted by a variety of factors rather than just the technology's performance or cost [9].

In the context of Malaysia, telehealth is becoming increasingly important in the health care system, especially with the older adult population expected to exceed 15% by 2030 [10]. This demographic shift will place added pressure on the health care system, particularly due to a rise in chronic diseases [11]. Malaysia's health care system offers accessible services but faces a workforce shortage to meet growing demands. With 2.4 physicians per 1000 people—fewer than those in Singapore, Japan, and Australia—Malaysia faces an aging population and staffing shortages leading to overcrowded public hospitals and strained health care capacity [11]. This has drastically burdened the health care system in Malaysia.

In response, telehealth offers a vital solution to address the growing imbalance between health care supply and demand as health care needs continue to rise [12]. The Malaysian government has been actively exploring technological solutions and launching telehealth initiatives to address the rising health care needs of its aging population. For instance, Malaysia's Ministry of Health initiated a teleconsultation at public hospitals to improve health care access and reduce congestion. Despite these efforts and the recent surge in telehealth-related studies, there remains a scarcity of research to investigate telehealth adoption in emerging economies, especially from the older adults' perspective [11,13]. Investigating telehealth adoption by older adults across different countries is essential, as varying levels of technological development and cultural contexts substantially influence their attitudes and behaviors [13,14].

Previous studies on technology adoption among older adults span various cultural contexts, revealing distinct factors influencing their behavior. In Canada, a study by Ahmed et al [15] found that >half of older adults adopted new technology for online social interactions. Despite having the knowledge to stay connected, they faced challenges like limited access and motivation. In a cross-cultural survey by Elimelech et al [16], older adults in Israel, France, and Spain exhibited different perceptions of technology use, emphasizing the need for culturally tailored adaptations. In Australia, Catapan et al [17] reported that patients had a high level of confidence and trust in the use of telehealth. Meanwhile, in China, Lin et al [18] showed that telehealth's ease of use and usefulness played a substantial role in affecting its adoption. In a developed country such as Singapore, Zhang et al [19] found that telehealth effectively supports the health-seeking behavior of older adults, challenging the belief that they resist technology and lack proficiency. Similarly, Haimi and Sergienko [20] found that telehealth uses among older adults in Israel remained elevated after the COVID-19 pandemic, indicating their ability to effectively learn and use digital health services.

However, in Malaysia, older adults may have different perceptions of telehealth. According to Ting et al [21], many older adults still prefer face-to-face interactions due to a cultural preference for personal consultations. Reservations about the impersonal nature of telehealth remain a substantial obstacle. However, the behavior of older adults toward telehealth in emerging economies, such as Malaysia, remains underexplored. There is a need for research focused on telehealth adoption within the Malaysian context to gain better grasp of this context.

Theory

Researchers have introduced various theoretical models to explain consumer behavior in the context of technology adoption. Well-established models for predicting technology acceptance among consumers include the theory of planned behavior by Ajzen [22], the technology acceptance model (TAM) by Davis et al [23], TAM2 by Venkatesh and Davis [24], the unified theory of acceptance and use of technology by Venkatesh et al [25], and TAM3 by Venkatesh and Bala [26], among others.

A growing body of research highlights the reliability and effectiveness of the TAM in explaining technology adoption. TAM, introduced by Davis et al [23], uses the concepts of perceived usefulness and perceived ease of use to elucidate how

technological factors influence a consumer's intent to adopt a specific technology. For the adoption of health-related technologies, the TAM and the unified theory of acceptance and use of technology have emerged as 2 prominent models, as demonstrated in studies by Harst et al [27], Heinsch et al [28], Rouidi et al [29], and Lin et al [18]. In the context of the aging population, a recent systematic literature review by Yap et al [30] confirmed that the TAM is the most widely used theory for explaining technology adoption among older adults.

Despite the widespread use of the TAM in the literature, previous studies argued that solely using the fundamental TAM to identify the consumer's technology adoption is insufficient [31,32]. Similarly, Attié and Meyer-Waarden [33] criticized that functional and utilitarian benefits, such as perceived ease of use, are insufficient to explain technology acceptance. Therefore, it can be found that previous studies have frequently extended the original TAM with additional variables or other theories to better reflect the technology's acceptance. For instance, Zhou et al [34] extended the original TAM by incorporating perceptions of medical service quality and information quality into the model for predicting telehealth acceptance among older adults. Telehealth acceptance, in turn, is influenced by older adults' perceptions of telehealth and their current behavioral intentions toward telehealth services. In addition, Rho et al [8] extended the TAM with perceived incentives, self-efficacy, and accessibility of patients' medical records, while Klingberg et al [35] included image, self-efficacy, voluntariness, compatibility, and anxiety in the TAM fundamental framework to enhance the explanatory power of the TAM. Moreover, integration of theories and perspectives has been performed as well, such as in the study by Tsai et al [36] that integrated the TAM with the status quo bias and technology anxiety concept to explain the telehealth intention. Therefore, it is suggested that there is a need for research to expand the TAM by including additional variables to provide a more comprehensive explanation and understanding regarding the telehealth intention of older adults.

Objectives

The objectives of this study are as follows:

- To investigate the coexistence and possible effects of TAM constructs, transition cost, and subjective well-being on telehealth service adoption among older adults
- To examine how inertia, availability, subjective well-being, and trust relate to the TAM's key antecedents

Methods

Research Model and Hypothesis Development

Overview

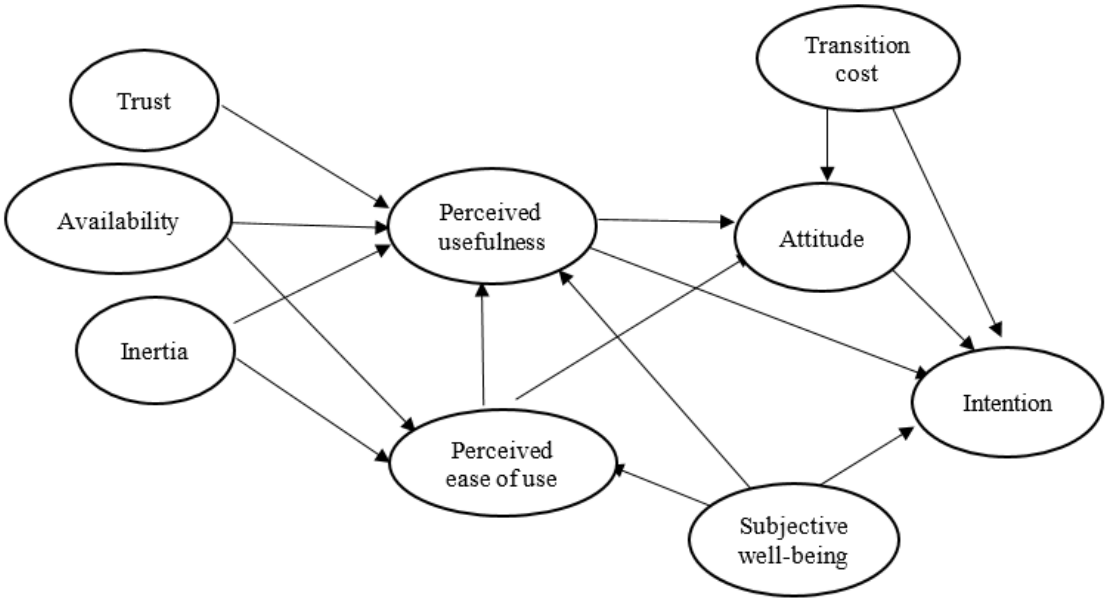
Telehealth is an effective and advanced alternative method for delivering health care services. In the context of telehealth adoption, the fundamental TAM might be insufficient to explain the telehealth adoption among older adults. In particular, the older adult population might often hold negative opinions about technology's inaccuracies that would influence their intention to use it [37]. Considering the evolving landscape of telehealth adoption among older adults in Malaysia and the limited existing literature, this study contributes to the expansion of the TAM. It does so by introducing and exploring a set of key factors aimed at comprehending the older adult population's willingness to embrace telehealth. In addition to attitude and perceived usefulness, this study proposes 2 additional constructs: transition cost and subjective well-being, to evaluate older adults' intention to use telehealth.

First, subjective well-being is included to address critiques of the traditional TAM. Recent studies [11,32,33] argue that the TAM's focus solely on utilitarian benefits is insufficient to fully explain consumer technology adoption. On the basis of transformative consumer studies and the uses and gratification theory [38], it is evident that beyond utilitarian benefits like usefulness and ease of use, consumers also seek affective elements, such as well-being when adopting a new technology [33]. The literature suggests enhancing the TAM by incorporating affective elements like subjective well-being to more accurately predict technology adoption [11]. Nevertheless, subjective well-being has received minimal attention in telehealth adoption studies, particularly from the perspective of the older adult population. Therefore, research proposes that subjective well-being be considered as one of the predictors of older adults' intention to use telehealth. If older adults perceive that telehealth can enhance their well-being, they are more likely to adopt it, as people naturally seek experiences that improve their overall quality of life [39,40].

The inclusion of transition cost as an extension to the TAM is grounded in the status quo bias theory, which posits that individuals tend to prefer maintaining their current routines over adopting change [41]. This is especially relevant for older adults, who often find transitions, such as shifting to telehealth, burdensome due to the perceived effort, time, and disruption of familiar health care practices like face-to-face consultations. Many older adults grew up in a time when technological innovations were not widespread, leading them to develop long-standing routines that provide comfort and predictability [42]. As a result, this study proposes that transition costs are particularly impactful for this group, as they are more likely to resist switching to telehealth in favor of maintaining their familiar health care practices.

Furthermore, the research incorporates availability, trust, inertia, and subjective well-being as factors influencing the TAM constructs. Subsequent sections of this study will provide a detailed analysis of the significance of these constructs within the unique context of this research. Figure 1 shows the research model of this study. The proposed hypotheses based on the developed research model are discussed in subsequent sections.

Figure 1. Research model.



Attitude

Attitude is an evaluation of effect, which refers to a person's positive or negative feelings regarding performing the respective behavior [43]. The impact of attitude on behavioral intention is a significant relationship in the theory of reasoned action, the theory of planned behavior, and the TAM. The connection between one's attitude and their intention signifies that individuals are more inclined to embrace technology if they hold a favorable perception of it, as noted by Davis et al [23]. The concept of attitude has played an important role in numerous studies seeking to gain deeper insights into consumers' willingness to adopt health care technologies. Previous research has consistently confirmed that a positive attitude toward health care–related technologies among consumers has a substantial impact on their intent to use such technologies, as evidenced by studies conducted by Park et al [44], Papa et al [45], Rajak and Shaw [46], and Ahn and Park [47].

Furthermore, Tsai et al [36] unveiled that the older adults' attitudes exert a positive influence on their intention to adopt telehealth. In addition, prior research has shown that attitude can serve as a mediator between beliefs and behavioral intent, as demonstrated by the work of Yang and Yoo [48]. Consequently, this study proposed the following hypothesis:

- Hypothesis 1: there is a positive association between attitude and the intention to adopt telehealth.

Perceived Ease of Use

Perceived ease of use is defined as the degree to which an older adult perceives that using telehealth technology would be free of effort. Perceived ease of use is one of the key constructs in the TAM. According to the TAM, perceived ease of use is related to attitudes and perceived usefulness of new technologies, which also influences intention [24]. Previous research has found that perceived ease of use influences consumers' attitudes and perceived usefulness toward new technologies [31,36,46,47]. For instance, Lazaro et al [49] confirmed that perceived ease of use positively affects the older adults' perceived usefulness and their attitude toward wearable health care technology. Hence, we proposed the following hypothesis:

- Hypothesis 2a: there is a positive association between perceived ease of use and perceived usefulness.
- Hypothesis 2b: there is a positive association between perceived ease of use and attitude.

Perceived Usefulness

Perceived usefulness is defined as the degree to which an individual feels that using a certain technology will improve their job performance [23]. Researchers have supported perceived usefulness as a crucial factor that predicts various types of technology adoption among older adults, such as the internet [50], social networking sites [51], health monitoring wearable technologies [52], telehealth [53], and automation technology [54]. Besides, previous studies revealed that perceived usefulness was highly relevant in predicting telehealth acceptance among the older adult population [34,55]. In addition, perceived usefulness influences consumers' attitudes toward new technologies [36,46,47]. Therefore, the following hypothesis is proposed:

- Hypothesis 3a: there is a positive association between perceived usefulness and attitude.
- Hypothesis 3b: there is a positive association between perceived usefulness and intention to adopt telehealth.

Transition Cost

According to Kim and Kankanhalli [56], transition costs refer to the user's perceived disutility that they incur when switching from the status quo to a new technology. When older adults consider using new technology, such as telehealth, the transition costs involved are essential. Transition costs that are uncertain might become a barrier and negatively influence a person's attitude [57]. In addition, Hsieh [58] revealed that the transition cost that might be incurred when using health-related technology

is one of the key concerns for technology adoption. An individual is likely to continue and remain with an existing system if the transition costs involved, such as effort and time, to learn to use a new technology, are deemed to be high. In the context of telehealth adoption, Tsai et al [36] claimed that transition costs negatively affect older adults' attitudes and hence affect their telehealth adoption. Therefore, if the transition cost to use a new technology is high, older adults will have a negative attitude toward it and not intend to use it. On the basis of the discussions, the following hypothesis is proposed:

- Hypothesis 4a: there is a negative association between transition cost and attitude.
- Hypothesis 4b: there is a negative association between transition cost and intention to adopt telehealth.

Subjective Well-Being

Subjective well-being refers to an individual's perception of an experience positively by using affective reactions and cognitive judgment instead of objective facts [59]. When adopting new technology, people always seek pleasurable experiences that enhance their well-being [39,40]. Previous studies have revealed that well-being influences consumer technology adoption [60]. Al-Jabri and Sohail [61] revealed that technology's characteristics contribute to a person's well-being. Similarly, well-being could act as a determinant that influences technology use [40]. Wu and Lu [62] have highlighted that positive emotions become a motivator for technology adoption when a person uses technology. In addition, findings show that well-being toward technology positively influences an individual's perceived ease of use, usefulness, and intention across all consumer adoption phrases [33,63,64]. Consumers will develop positive emotions toward technology when they perceive that using it will enhance their well-being. Therefore, positive emotions will positively influence the perceived benefits, such as perceived usefulness and ease of using the technology [33].

Therefore, this study proposes that older adults perceive that telehealth will enhance their well-being, which consequently will influence their perceived ease of use, usefulness, and intention to use telehealth. Therefore, we hypothesize the following:

- Hypothesis 5a: there is a positive association between well-being and perceived ease of use.
- Hypothesis 5b: there is a positive association between well-being and perceived usefulness.
- Hypothesis 5c: there is a positive association between well-being and intention to adopt telehealth.

Inertia

Inertia refers to the degree of a person's willingness to continue using traditional physical products despite knowing that better options are available [41]. Even when better alternatives or switching incentives are available, consumers remain attached to and steadfast in their use of existing technologies [65]. Hence, the greater an individual's attachment toward a thing that he or she is familiar with, the less inclined they are to explore new experiences. In addition, Bem [66] and Petty and Cacioppo [67] claimed that people usually depend on their prior behavior and hence fail to recognize a new technology's advantages. In line with this assumption, older adults always seek to maintain their existing internal and external structures when making adaptive decisions, as they would prefer to continue to engage in similar activities or behaviors as they did throughout their previous experiences [68]. The older adult population grew up in an era when technological innovation was not commonly used. As a result of their early experiences, they might have long been accustomed to receiving health care services physically at hospitals or clinics. Older adults might perceive telehealth as not useful and not easy to operate as they prefer to continue to engage in the behavior they are more familiar with to minimize feelings of anxiety. It is hypothesized that inertia will negatively influence the older adults' behavioral perceptions of a new technology and hence create lower inclinations to use new technologies. Individuals who have high inertia tend to reduce the variety of technologies that are available to them and rely on prior behavior to influence their perceptions and intentions [36]. Therefore, the following hypothesis is proposed:

- Hypothesis 6a: there is a negative association between inertia and perceived ease of use.
- Hypothesis 6b: there is a negative association between inertia and perceived usefulness.

Availability

On the basis of the study by Venkatesh [69], availability refers to the extent to which consumers perceive that they can obtain a technological service or product without barriers, along with the presence of organizational support to help them overcome any challenges in using the technology. A person's control belief about the availability of organizational resources and support structures to enable technology use is related to facilitating conditions [26]. Many previous studies have proved the positive impact of facilitating conditions on technology adoption [70,71]. In the telehealth adoption context, the determinant associated with external control might involve the availability of manufacturer's assistance, where the firms provide consumers assistance to overcome the difficulties of using a new technology. Telehealth technologies can make medical resources available to health care professionals, caregivers, and older adults at any time and from any location, allowing a considerable improvement in patient health care. Wu et al [72] revealed that availability positively influences the perceived usefulness of the telehealth care technology. In addition, previous studies also provided support on the impact of availability in the context of telehealth adoption [73,74]. Similarly, a recent study by Tsai et al [36] revealed that availability is an important predictor in determining perceived ease of use and usefulness of telehealth among the older adult population. According to the existing evidence, this study hypothesizes that availability would increase the older adults' perceived ease of use and usefulness of telehealth.

- Hypothesis 7a: there is a positive association between availability and perceived ease of use.
- Hypothesis 7b: there is a positive association between availability and perceived usefulness.

Trust

Trust is evidently a crucial determinant in health care adoption. On the basis of the study by Gefen et al [75], trust refers to a sense of confidence in the trustworthiness and integrity of the other party. When it comes to technological adoption, trust is essential, especially when the technology is relatively new and might involve risks and uncertainties for the older adult population. Despite trust not being included in the original TAM, it has been incorporated into several of the study contexts. Many existing studies demonstrate that trust has a strong positive influence on technology adoption [45,76]. Previous studies also revealed that trust plays an important role in predicting adoption of health-related technologies (eg, telehealth) [76,77].

In the context of this study, telehealth can be a difficult and complex task, especially for older adults, as its use requires a good understanding of devices to communicate effectively with the health care providers [78,79]. Hence, people might lack trust in using telehealth due to the risks incurred, including unclear regulatory authority in place to deal with issues like confidentiality, misconduct, and liability in telehealth. Older adults, who used to obtain services physically, might have trust issues with telehealth's ability to replace in-person consultations and physical health assessments. Particularly in the IT setting, Li et al [80] revealed that trust is crucial, as people must overcome the perceived risk before technology adoption. Previous research has established the importance of trust in determining health-related technology adoption [46]. For instance, Catapan et al [17], Chew et al [81], and Orrange et al [82] showed that trust substantially affects telehealth adoption. Hence, this study hypothesizes that trust would increase older adults' perceived usefulness of telehealth.

- Hypothesis 8: there is a positive association between trust and perceived usefulness.

Research Instrument Development

The measurement scales were adjusted in accordance with existing literature and tailored to suit our research context. The questionnaire was divided into 2 sections. The first part consisted of the 11 constructs used in this study: attitude, availability, transition cost, perceived ease of use, perceived usefulness, inertia, trust, subjective well-being, and intention to adopt. The second part comprised a survey focusing on demographic characteristics, such as gender and ethnicity. An overview of the instrument is provided in Multimedia Appendix 1 [32,36,83,84].

The scales for attitude, availability, transition cost, perceived ease of use, perceived usefulness, and inertia were taken from the study of Tsai et al [36] and Zhang and Zaman [83]. The measures for subjective well-being were based on the work of Yap et al [32]. Trust and intention were measured according to the work of Wu et al [84]. In total, 39 items were assessed using a 5-point Likert scale, which ranged from *completely disagree* (score=1) to *completely agree* (score=5).

To ensure the questionnaire's reliability, a pilot survey was carried out involving 10 older adults. This pretest, which involved contacting 10 older adults before conducting the web-based survey, was conducted to validate the instrument. On the basis of the feedback received, minor adjustments were made to the questionnaire to improve its effectiveness.

Research Sample and Data Collection Procedure

This study focused on the factors influencing the intention to adopt telehealth among adults aged ≥60 years in Malaysia. This demographic was chosen as research participants due to their increased susceptibility to chronic diseases, as outlined by Tsai et al [36]. Chronic diseases can affect individuals of all ages, but the risk escalates as people advance in age, justifying the selection of this population for our research.

In our data collection process, we used a survey approach to investigate the intent of older adults to adopt telehealth services. Recognizing the challenges in directly reaching this demographic, we engaged students as intermediaries to connect with their family members who were older. This method combines elements of convenience and snowball sampling, as students were encouraged to distribute surveys within their social networks, primarily targeting their older family members. We encouraged student intermediaries to recruit participants from diverse geographic regions and socioeconomic backgrounds to enhance sample diversity where possible.

The rationale behind using student intermediaries was 2-fold: first, it provided a practical and effective means of reaching older adult participants, a group that may not be as digitally connected or comfortable with technology. The student intermediaries, who were trained to administer the survey, acted as trusted conduits, facilitating communication and engagement with older adult participants in a familiar, trustworthy, and less intimidating environment. This helped to overcome potential barriers related to accessibility and comprehension of the survey, ensuring that the older adult participants felt comfortable and supported throughout the process. Hence, it was essential for the student intermediaries to have a thorough understanding of both the study's purpose and the questionnaire to effectively guide their older family members. To ensure this, students underwent a comprehensive briefing before the survey, equipping them with a solid grasp of the questionnaire and its objectives.

Data Analysis

This study conducted data analysis using partial least squares structural equation modeling (PLS-SEM). We used SmartPLS 3.2.8 [85] for PLS-SEM, as it is well-suited for analyzing measurement and structural models without the need for normality assumptions. This is particularly useful because survey research data are often not normally distributed [86]. Furthermore, PLS-SEM offers a higher explanatory power compared to covariance-based structural equation modeling.

Ethical Considerations

This study received formal approval from the Research Ethics Committee of Multimedia University (EA2882021). Informed written consent was obtained from all survey respondents before participation. Respondents were required to read the ethical statement at the top of the survey form and proceed only if they agreed to participate. All collected data are treated with the utmost confidentiality, ensuring anonymity and used solely for

research purposes. Additionally, consent to publish was obtained from all participants.

Results

Respondent Characteristics

The survey was conducted from March 1, 2022, to August 31, 2022, resulting in 125 received samples, of which 119 were considered valid. While this approach offers advantages for reaching a challenging-to-access population, researchers should remain vigilant about potential biases arising from the familial and social connections of the student intermediaries. In the sample, 52 (43.7%) of the 119 respondents were women, and 67 (56.3%) were men. Additional demographic details are provided in Multimedia Appendix 2.

Reliability and Validity Tests

The evaluation of the measurement model involved assessing reliability and validity. In Table 1, all variables exhibited factor loadings >0.7 [87], and Cronbach α exceeded 0.7, signifying a high level of reliability. Both the composite reliability and average variance extracted surpassed 0.7 and 0.5, respectively, indicating strong convergent validity [88]. In the second step of the analysis, we evaluated discriminant validity using the heterotrait-monotrait (HTMT) ratio criterion, as suggested by Henseler et al [89] and Franke and Sarstedt [90]. As indicated in Table 2, the results of the HTMT criterion demonstrate that all HTMT values fall below the threshold of 0.85 for the more stringent criterion. This indicated that the respondents recognized the distinctiveness of all the constructs. The combination of these 2 validity tests affirms that the measurement items exhibited both validity and reliability.

Table 1. Measurement model and cross-validated redundancy.

Constructs and items	Loadings	Composite reliability	Average variance extracted
Attitude		0.896	0.743
ATT[a]1	0.833		
ATT2	0.850		
ATT3	0.900		
Perceived ease of use		0.963	0.897
EOU[b]1	0.938		
EOU2	0.954		
EOU3	0.948		
Perceived usefulness		0.928	0.762
USE[c]1	0.841		
USE2	0.857		
USE3	0.902		
USE4	0.890		
Transition cost		0.953	0.872
COST[d]1	0.920		
COST2	0.939		
COST3	0.942		
Inertia		0.889	0.728
INE[e]1	0.889		
INE2	0.811		
INE3	0.858		
Availability		0.886	0.723
AVAI[f]1	0.862		
AVAI2	0.888		
AVAI3	0.798		
Trust		0.900	0.751
TRUST[g]1	0.886		
TRUST2	0.808		
TRSUT3	0.903		
Subjective well-being		0.959	0.885
SWB[h]1	0.934		
SWB2	0.948		
SWB3	0.941		
Intention		0.932	0.821
INT[i]1	0.911		
INT2	0.885		
INT3	0.923		

[a]ATT: attitude.

[b]EOU: perceived ease of use.

[c]USE: perceived usefulness.

[d]COST: transition cost.

[e]INE: inertia.
[f]AVA: availability.
[g]TRUST: trust.
[h]SWB: subjective well-being.
[i]INT: intention.

Table 2. Discriminant validity (heterotrait-monotrait 0.85 criterion).

	ATT[a]	EOU[b]	USE[c]	COST[d]	INE[e]	AVA[f]	TRUST[g]	SWB[h]	INT[i]
ATT	—[j]	—	—	—	—	—	—	—	—
EOU	0.677	—	—	—	—	—	—	—	—
USE	0.769	0.719	—	—	—	—	—	—	—
COST	0.549	0.623	0.498	—	—	—	—	—	—
INE	0.390	0.427	0.261	0.523	—	—	—	—	—
AVA	0.665	0.654	0.781	0.453	0.140	—	—	—	—
TRUST	0.688	0.530	0.683	0.389	0.210	0.527	—	—	—
SWB	0.755	0.577	0.715	0.170	0.374	0.545	0.753	—	—
INT	0.753	0.674	0.701	0.570	0.432	0.579	0.677	0.818	—

[a]ATT: attitude.
[b]EOU: perceived ease of use.
[c]USE: perceived usefulness.
[d]COST: transition cost.
[e]INE: inertia.
[f]AVA: availability.
[g]TRUST: trust.
[h]SWB: subjective well-being.
[i]INT: intention.
[j]Not applicable.

Results of the Structural Model

Following the guidance of Hair et al [91], we reported the path coefficients, SEs, 1-tailed *t* test values, and *P* values using a bootstrapping procedure with 5000 resamples [92]. This large resample size ensures result stability, according to Hair et al [91] and Hair and Alamer [93]. In addition, Hahn and Ang [94] emphasized that *P* values alone were insufficient for hypothesis significance and recommended combining *P* values, effect sizes, and bias-corrected interval for a more comprehensive evaluation. Table 3 presents the evaluation of the hypotheses. Moreover, the results of bias-corrected interval, effect sizes, and variance inflation factor are presented in Multimedia Appendix 3.

Table 3. Hypothesis testing direct effects[a].

Hypothesis	Relationship	Standard β	SD	*t* test value	*P* value
Hypothesis 1	Attitude → intention	0.242	0.058	4.168	<.001
Hypothesis 2a	Perceived ease of use → perceived usefulness	0.271	0.048	5.693	<.001
Hypothesis 2b	Perceived ease of use → attitude	0.209	0.056	3.734	<.001
Hypothesis 3a	Perceived usefulness → attitude	0.456	0.054	8.450	<.001
Hypothesis 3b	Perceived usefulness → intention	0.100	0.050	1.997	.02
Hypothesis 4a	Transition cost → attitude	−0.153	0.055	2.779	.003
Hypothesis 4b	Transition cost → intention	−0.163	0.037	4.357	<.001
Hypothesis 5a	Subjective well-being → perceived ease of use	0.260	0.053	4.927	<.001
Hypothesis 5b	Subjective well-being → perceived usefulness	0.261	0.057	4.531	<.001
Hypothesis 5c	Subjective well-being → intention	0.448	0.047	9.582	<.001
Hypothesis 6a	Inertia → perceived ease of use	−0.246	0.041	5.972	<.001
Hypothesis 6b	Inertia → perceived usefulness	0.024	0.034	0.778	.22
Hypothesis 7a	Availability → ease of use	0.420	0.048	8.725	<.001
Hypothesis 7b	Availability → perceived usefulness	0.323	0.047	6.818	<.001
Hypothesis 8	Trust → perceived usefulness	0.156	0.060	2.620	.004

[a]We used 95% CI with a bootstrapping of 5000.

First, we examined the influence of the 4 predictors on intention. Attitude (β=0.242; *P*<.001), perceived usefulness (β=0.100, *P*=.02), transition cost (β=−0.163; *P*<.001), and subjective well-being (β=0.448; *P*<.001) were all associated with intention. This means that hypotheses 1, 3b, 4b, and 5c were supported.

Second, the impact of the 3 predictors on attitude was investigated. Perceived ease of use (β=0.209; *P*<.001), perceived usefulness (β=0.456; *P*<.001), and transition cost (β=−0.153; *P*<.001) were linked to attitude, affirming support for hypotheses 2b, 3a, and 4a.

Third, perceived usefulness was assessed in relation to 5 predictors. Perceived ease of use (β=0.271; *P*<.001), availability (β=0.323; *P*<.001), subjective well-being (β=0.261; *P*<.001), and trust (β=0.156; *P*<.001) exhibited positive associations, while inertia (β=.024, *P*=.22) showed no significant relationship with perceived usefulness. Consequently, hypotheses 2a, 5b, 7b, and 8 were supported, while hypothesis 6b was not.

Finally, we examined the impact of availability (β=.420; *P*<.001) and subjective well-being (β=0.260; *P*<.001) on perceived ease of use, finding that both were positively related, while inertia (β=−0.246; *P*<.001) was negatively related. This supported hypotheses 5a, 6a, and 7a.

Consistent with the results of previous hypothesis testing, the bias-corrected 95% CIs for all hypotheses (except 6a) confirmed that they did not encompass 0, indicating support for these hypotheses (Multimedia Appendix 3). Furthermore, the variances for intention to adopt (R^2=0.650), attitude (R^2=0.502), perceived usefulness (R^2=.654), and perceived ease of use (R^2=0.475) were generally above 33%. This suggests the model possesses a moderate predictive capacity.

To assess the predictive relevance of the model, Shmueli et al [95] suggested the use of PLSpredict, a holdout sample-based method that produces case-level predictions at either the item or construct level. We used PLSpredict with a 10-fold procedure to assess predictive capability. As demonstrated in Table 4, all the errors associated with the PLS model were lower compared to the linear model, indicating the robust predictive power of our model. Therefore, we can confidently assert that our model exhibits strong predictive capacity.

Table 4. PLSpredict results.

Items for the construct intention	PLS[a]	LM[b]	PLS-LM
INT1[c]	0.734	0.735	−0.001
INT2	0.856	0.866	−0.010
INT3	0.767	0.769	−0.002

[a]PLS: partial least squares.

[b]LM: linear model.

[c]INT: intention.

Results of the Importance-Performance Map Analysis

The importance-performance map analysis (IPMA) of older adults' intention to adopt telehealth services aimed to expand upon the results of the structural model by evaluating the performance of each variable. As mentioned by Hair et al [91], areas requiring management attention are those with high importance but poor performance on a specific endogenous latent variable. In our research, we assessed the impact of latent exogenous factors on the endogenous variable (ie, intention to adopt) in terms of their significance and performance. The results of this analysis are presented in Table 5.

Table 5. Importance-performance map analysis.

Constructs	Performance	Importance
Attitude	63.245	0.241
Perceived ease of use	60.779	0.108
Perceived usefulness	64.078	0.210
Transition cost	52.758	0.200
Inertia	69.450	0.021
Availability	70.735	0.113
Trust	64.529	0.033
Subjective well-being	53.330	0.532

The IPMA results showed that the most important factor was subjective well-being (0.532), followed by attitude (0.241), perceived usefulness (0.210), transition cost (0.200), availability (0.113), perceived ease of use (0.108), trust (0.033), and inertia (0.021).

On the basis of performance, availability (70.735) was the highest, followed by inertia (69.45), trust (64.53), perceived useful (64.078), attitude (63.22), ease of use (60.779), subjective well-being (53.33), and transition cost (52.758). Therefore, it is evident that subjective well-being, attitude, transition cost, and perceived usefulness play crucial roles in influencing the intention of older adults to adopt telehealth, as these constructs exhibit relatively higher total effects (importance) compared to other factors in the model. However, the performance of well-being, attitude, transition cost, and usefulness were relatively lower compared to other factors like availability, inertia, and trust.

In summary, to enhance the intention to adopt telehealth among older adults, managerial efforts should be primarily focused on addressing and emphasizing subjective well-being, while continuing to support and maintain positive attitudes and perceived usefulness. On the other hand, perceived ease of use should be given lower priority.

Discussion

Principal Findings

This study aims to enhance the existing TAM for the adoption of telehealth services among older adults in the Malaysian context. Apart from attitude and perceived usefulness, this study proposed 2 additional constructs, namely subjective well-being and transition cost, for assessing the intention of older adults to use telehealth. Furthermore, this study introduced availability, trust, inertia, and subjective well-being as antecedents of TAM constructs. The study found that subjective well-being is the most important factor in telehealth adoption, followed by attitude, transition cost, and perceived usefulness. Perceived ease of use, perceived usefulness, and transition cost substantially affected attitude. Perceived ease of use, availability, subjective well-being, and trust positively influenced perceived usefulness, while inertia did not. In addition, availability and subjective well-being were positively related to perceived ease of use, with inertia having a negative impact. IPMA results showed that subjective well-being was the most crucial factor for older adult users, while availability had the highest performance in telehealth services.

As expected, this study found that attitude positively affects intention, which is in line with the existing health care–related studies, such as those by Park et al [44], Papa et al [45], Rajak and Shaw [46], and Ahn and Park [47]. As per the results, the key TAM construct, perceived ease of use, positively influences older adults' perceived usefulness and attitude toward the telehealth system. These outcomes corroborate the findings of Rajak and Shaw [46] and Ahn and Park [47], which indicate that older adults will have a positive attitude toward telehealth and perceive it to be useful to them if telehealth is easy to use. On the other hand, perceived usefulness also positively influences older adults' attitude and intention toward telehealth. Specifically, perceived usefulness has a significantly larger influence on attitudes regarding adopting telehealth than perceived ease of use (0.456 vs 0.209). The rationale behind these findings may be due to the relevance of using telehealth for managing older adults' health. Hence, the usefulness of telehealth is the priority as compared to its ease of use. In other words, older adults are more likely to use telehealth services if they can provide useful features to them, such as improving their quality of life and offering better health care services.

Consistent with previous findings [36,57,58], these findings show that transition costs are driving forces that have a negative impact on the older adults' attitude and intention to use telehealth. According to the findings, older adults do not intend to use telehealth and will continue using the traditional method of obtaining health care services if they believe the time and effort required to learn telehealth is too high.

Besides, inertia was found to negatively influence the perceived ease of use but had no impact on the perceived usefulness. The significance of inertia's influence on perceived ease of use stems from the fact that ease of use is directly linked to how simple or user-friendly telehealth appears. When older adults experience inertia, they tend to resist adopting telehealth because of the additional effort required to learn and adapt to new technologies. This resistance (inertia) makes new systems seem more complex, thereby reducing the perceived ease of use of telehealth among older adults. However, the significance of inertia on perceived ease of use contradicts the findings of Tsai et al [36], where inertia had no effect on ease of use. This discrepancy may stem from the fact that the majority of respondents in the study by Tsai et al [36] were younger (>40 years) with higher digital literacy, while our study focuses on those aged ≥60 years. In contrast, the insignificant impact of inertia on perceived usefulness in this study is consistent with the findings of Tsai et al [36]. This is because perceived usefulness is more about the functional benefits the user gains from the technology. Hence, even if older adults perceive the system to be hard to use due to inertia, they might still acknowledge that the system provides value or could be beneficial in terms of performance and efficiency. This can explain why inertia does not affect perceived usefulness to the same extent.

Availability showed a positive influence on ease of use and perceived usefulness, which is consistent with existing studies, such as those by Chang et al [74] and Tsai et al [36]. However, these findings partially conflict with the findings of Wu et al [72], which revealed that availability only influences the perceived usefulness and has no impact on the perceived ease of use. This revealed different perspectives between patients and hospital professionals toward the importance of availability in determining telehealth adoption. From the older adult users' perspective, the study's findings revealed that the availability of assistance, such as providing responses to assist them in overcoming obstacles when using telehealth, is very important, which would affect their perceived ease of use and usefulness of telehealth.

This study revealed that well-being positively influences perceived ease of use, perceived usefulness, and intention toward telehealth. These outcomes are expected as people always seek pleasurable experiences that can enhance their well-being when adopting new technology [39,40]. These findings are also consistent with previous results, which showed that an individual's perceived well-being toward technology will positively influence their perceived ease of use, usefulness, and intention [32,33,63,64].

Furthermore, this study also revealed that older adults' trust positively influences the perceived usefulness and attitude of telehealth. This finding is predictable because most older adults prefer to receive services in-person and may have trust issues with telehealth's ability to replace in-person consultations and physical health assessments, so they may perceive high risks and lower trust toward telehealth. The study validates the findings of Rajak and Shaw [46], who revealed the crucial role of trust in the health care technology adoption among older adults.

Conclusions and Implications

Given the rapid growth of the older adult health care industry, improving health care–related services is a promising opportunity. Hence, this study was conducted to investigate the factors influencing the intention of older adults to use telehealth services by extending the TAM model. The model has been stretched to incorporate factors like availability, transition cost, trust, inertia, and well-being. From the literature, it was observed that a considerable number of studies have been conducted on telehealth. However, there is no evidence in the literature of incorporating these factors and further assessing the model in the Malaysian context. In addition, although several papers have adopted the TAM to investigate telehealth adoption, there are limited studies investigating the formation of the TAM key constructs (perceived ease of use and perceived usefulness) in depth from the older adults' perspective. With the introduction of these constructs into the TAM model, this study can contribute to the literature by providing a better understanding of factors affecting older adults' intention to use telehealth in the Malaysian context.

This study has several practical implications. First, the positive impact of attitude toward older adults' intention on telehealth use provides valuable implications to the health care centers and managers. They should create an environment that can ensure older adults have a positive attitude toward telehealth. On the basis of the findings, positive attitude toward telehealth can be enhanced in several ways, for example, when the older adults perceive it to be useful and easy to use. Hence, the telehealth developer and health care personnel should carefully assess telehealth's usefulness in assisting older adults' health care before it is introduced to them.

In addition, telehealth developers should also design an older adult–friendly interface that is easier for them to navigate and understand, as the older adults are not as digitally literate as the younger generation. For example, a case study by Pires et al [96] emphasizes the importance of simplicity and ease of use in successfully implementing the VITASENIOR-MT telehealth system for older adult users. By using television as the primary interface, the telehealth system became more accessible and easier by integrating familiar technology into the homes of older adults. Health professionals provided valuable feedback on design and usability, enhancing the system's effectiveness as they remotely monitored patients. This case exemplifies the best practices in telehealth design, highlighting the significance of user-centered development that prioritizes ease of use for older adults. In addition, health professionals are encouraged to actively engage in the development process to further improve the system's effectiveness. A clear step-by-step video tutorial, easy interactive dialogue, and straightforward click-through procedures for making health care appointments can be introduced. Moreover, the government or relevant authorities can provide older adults with training on how to use telehealth services to improve their skills and confidence in accessing online services.

The results from this study revealed that if the transition cost required when using telehealth is perceived as high, it will cause people to have a negative attitude and intention toward

telehealth. Furthermore, results also indicate that older adults tend to keep using their habitual ways to obtain health care services, which is known as inertia in this study. This inertia negatively affects their perceived ease of use and usefulness of telehealth. The rationale behind these results might be related to the fact that most commonly available smartphone devices, such as tablet computers and smartphones, as well as mobile internet subscription plans, are generally sold at a high price. Therefore, older adult mobile users only seek to access the internet through free Wi-Fi networks due to the high cost of using mobile devices. Given these considerations, it is recommended that telecommunications companies collaborate with government agencies to launch programs that offer special low prices, rebates, or government-funded subsidies for older adults when they purchase a mobile device and sign up for mobile internet packages. For instance, telecommunication companies in Taiwan have effectively held a promotional campaign for older adults to access the internet for free for a certain period [97]. Through a short, free trial, older adults' increasing internet accessibility is an effective way to encourage them to learn and become familiar with mobile apps and to reduce their inertia by providing a greater understanding of the mobile services' value and usefulness, particularly the telehealth services.

In addition to utilitarian-oriented benefits like usefulness and ease of use, which have been frequently discussed in earlier literature, policy makers and managers must also highlight the affective components like well-being [33]. Surprisingly, the result revealed that well-being is the most important factor that affects older adults' intention for telehealth use. Hence, practitioners should pay more attention to developing a sense of well-being enhancement if older adults use telehealth services. For instance, relevant authorities can create an online community of innovative users where the older adults can share their pleasant, positive, and healthy experiences to highlight the well-being provided by telehealth, and hence indirectly generate positive word of mouth that motivates the potential users to use telehealth. On the other hand, findings also revealed that availability positively influences older adults' perceived ease of use and usefulness of telehealth. Therefore, technology developers, manufacturers, or organizations should always be prepared to respond and assist the older adult in overcoming challenges and obstacles when using telehealth services, particularly during the early stages of adoption.

Furthermore, the study's findings also revealed that older adults' trust positively affects the perceived usefulness. Incorporating trust as a key determinant of perceived usefulness is particularly important for older adults adopting telehealth. The literature suggests that trust in telehealth can be fostered through mechanisms, such as confidence in the competence of health care providers [17,81,82] and the reliability of the information provided [98]. In practice, trust can be strengthened by implementing ongoing professional development and telehealth-specific training for health care providers. This can ensure they remain updated on best practices and patient care, thus boosting patient confidence in their competence. In addition, regularly updating health information and cross-referencing it with evidence-based guidelines reinforces the reliability of the information provided on telehealth platforms.

Moreover, government authorities and telehealth platform developers play a critical role in improving trust in using telehealth services. As a result, it is suggested that government agencies implement clear and detailed support policies for telehealth services. For instance, the government should strengthen the laws to ensure that the platform developers ensure the confidentiality of patient data and data security while health care institutions and staff are qualified. The introduction of national policies shows the government's support for telehealth services, which could increase the public's trust toward them [99]. In addition, managers should assist in the development of telehealth platforms by implementing and promoting features and guidelines on patient data confidentiality and privacy protection to establish customer trust. Furthermore, the use of secondary data and security policies should be communicated and advertised by companies to increase trust. In addition, the government should also provide sufficient budget allocation and training for health care professionals to use telehealth services so they can provide better telehealth services for patients and, in turn, increase their trust in telehealth services.

Limitations and Suggestions for Future Research

Like all research endeavors, this study is not immune to limitations. The first constraint becomes evident through the limited number of determinants examined concerning the intention to use telehealth services. We recognize that numerous potential factors remain unexplored, particularly within the realm of customization.

Second, our research adopts a cross-sectional survey design approach. Given the rapid evolution of technology and the increasing familiarity of consumers with telehealth, this approach may impact their evolving perceptions of telehealth. In addition, consumers' lifestyles are dynamic, undergoing continuous changes across different life stages. Consequently, a longitudinal study would yield valuable insights into the sustained use of telehealth services. In addition, future research could explore how different types of telehealth services (eg, video consultations vs remote monitoring) specifically impact adoption rates among older adults.

This study used convenient and snowball sampling to reach a hard-to-access population. Although the data represent various ethnicities and regions across Malaysia, the uneven distribution of responses may still introduce a certain level of bias. To enhance representativeness, we recommend using random or stratified sampling methods in future research.

Another limitation is the potential for self-report biases, as participants may have provided socially desirable responses rather than their true thoughts or behaviors. While we mitigated this by ensuring survey anonymity, self-report biases may still persist and should be considered when interpreting the findings. Future research should incorporate objective measures, such as tracking actual telehealth use data, to complement self-reported data and provide a more accurate assessment.

Although this study meets the sample size requirements based on power analysis using G-Power, it may not fully capture the

diversity of the older adult population, potentially limiting the generalizability of the findings. Caution is advised when applying these results to the broader older adult population. Future studies should include larger, more diverse samples across different socioeconomic backgrounds, geographic locations, and health conditions to strengthen the robustness and applicability of the findings.

Acknowledgments

The authors acknowledge financial support from the Ministry of Higher Education, Malaysia (FRGS/1/2019/SS06/MMU/02/3), and Multimedia University (MMUI/220112).

Authors' Contributions

SHT and YYY: conceptualization, methodology, writing, analysis and interpretation of data, and software. SHT: funding acquisition. SKT: review and editing and visualization. CKW: acquisition of data. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of the instrument.
[DOCX File , 21 KB - aging_v8i1e60936_app1.docx]

Multimedia Appendix 2
Respondents' information (N=119).
[DOCX File , 15 KB - aging_v8i1e60936_app2.docx]

Multimedia Appendix 3
Bias-corrected interval, f2, and variance inflation factor.
[DOCX File , 17 KB - aging_v8i1e60936_app3.docx]

References

1. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
2. Khosravi P, Ghapanchi AH. Investigating the effectiveness of technologies applied to assist seniors: a systematic literature review. Int J Med Inform 2016 Jan;85(1):17-26. [doi: 10.1016/j.ijmedinf.2015.05.014] [Medline: 26216463]
3. Pinto A, Almeida JP, Pinto S, Pereira J, Oliveira AG, de Carvalho M. Home telemonitoring of non-invasive ventilation decreases healthcare utilisation in a prospective controlled trial of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 2010 Nov;81(11):1238-1242. [doi: 10.1136/jnnp.2010.206680] [Medline: 20826878]
4. Gellis ZD, Kenaley B, McGinty J, Bardelli E, Davitt J, Ten Have T. Outcomes of a telehealth intervention for homebound older adults with heart or chronic respiratory failure: a randomized controlled trial. Gerontologist 2012 Aug;52(4):541-552. [doi: 10.1093/geront/gnr134] [Medline: 22241810]
5. Chou CC, Chang CP, Lee TT, Chou HF, Mills ME. Technology acceptance and quality of life of the elderly in a telecare program. Comput Inform Nurs 2013 Jul;31(7):335-342. [doi: 10.1097/NXN.0b013e318295e5ce] [Medline: 23728446]
6. Nikou S, Agahari W, Keijzer-Broers W, de Reuver M. Digital healthcare technology adoption by elderly people: a capability approach model. Telemat Inform 2020 Oct;53:101315. [doi: 10.1016/j.tele.2019.101315]
7. Hanson VL. Influencing technology adoption by older adults. Interact Comput 2010 Nov;22(6):502-509. [doi: 10.1016/j.intcom.2010.09.001]
8. Rho MJ, Choi IY, Lee J. Predictive factors of telemedicine service acceptance and behavioral intention of physicians. Int J Med Inform 2014 Aug;83(8):559-571. [doi: 10.1016/j.ijmedinf.2014.05.005] [Medline: 24961820]
9. Lee C, Coughlin JF. PERSPECTIVE: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Product Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
10. Current population estimates, Malaysia, 2023. Ministry of Economy, Department of Statistics Malaysia. 2023. URL: https://www.dosm.gov.my/portal-main/release-content/current-population-estimates-malaysia----2023 [accessed 2024-04-12]
11. Tan SH, Wong CK, Yap YY, Tan SK. Factors influencing telemedicine adoption among physicians in the Malaysian healthcare system: a revisit. Digit Health 2024 Jun 06;10:20552076241257050 [FREE Full text] [doi: 10.1177/20552076241257050] [Medline: 38854922]

12. Lawrence K, Nov O, Mann D, Mandal S, Iturrate E, Wiesenfeld B. The impact of telemedicine on physicians' after-hours electronic health record "work outside work" during the COVID-19 pandemic: retrospective cohort study. JMIR Med Inform 2022 Jul 28;10(7):e34826 [FREE Full text] [doi: 10.2196/34826] [Medline: 35749661]
13. Teh PL, Kwok AO, Cheong WL, Lee S. Insights into the use of a digital healthy aging coach (AGATHA) for older adults from Malaysia: app engagement, usability, and impact study. JMIR Form Res 2024 May 21;8:e54101 [FREE Full text] [doi: 10.2196/54101] [Medline: 38772022]
14. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506 [FREE Full text] [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
15. Ahmed MH, Awol SM, Kanfe SG, Hailegebreal S, Debele GR, Dube GN, et al. Willingness to use telemedicine during COVID-19 among health professionals in a low income country. Inform Med Unlocked 2021;27:100783 [FREE Full text] [doi: 10.1016/j.imu.2021.100783] [Medline: 34778509]
16. Elimelech OC, Ferrante S, Josman N, Meyer S, Lunardini F, Gómez-Raja J, et al. Technology use characteristics among older adults during the COVID-19 pandemic: a cross-cultural survey. Technol Soc 2022 Nov;71:102080 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102080] [Medline: 35991080]
17. Catapan SD, Haydon HM, Hickman IJ, Webb L, Isbel N, Johnson D, et al. Trust and confidence in using telehealth in people with chronic kidney disease: a cross-sectional study. J Telemed Telecare 2023 Nov 26;29(10_suppl):16S-23S. [doi: 10.1177/1357633x231202275]
18. Lin Y, Xu X, Liu Y, Alias H, Hu Z, Wong LP. Perception and acceptance of telemedicine use in health care among the general public in China: web-based cross-sectional survey. J Med Internet Res 2024 Jul 16;26:e53497 [FREE Full text] [doi: 10.2196/53497] [Medline: 39012687]
19. Zhang Y, Lee EW, Teo WP. Health-seeking behavior and its associated technology use: interview study among community-dwelling older adults. JMIR Aging 2023 May 04;6:e43709 [FREE Full text] [doi: 10.2196/43709] [Medline: 36996003]
20. Haimi M, Sergienko R. Adoption and use of telemedicine and digital health services among older adults in light of the COVID-19 pandemic: repeated cross-sectional analysis. JMIR Aging 2024 Apr 24;7:e52317 [FREE Full text] [doi: 10.2196/52317] [Medline: 38656768]
21. Ting CY, Abdul Halim NH, Ling JN, Tiong IK, Ahmad Shauki NI, Lee YF, et al. The use of a multi-disciplinary geriatric telemedicine service (TELEG) and its acceptance at a tertiary care centre in Malaysia. BMC Geriatr 2024 Feb 05;24(1):133 [FREE Full text] [doi: 10.1186/s12877-024-04676-0] [Medline: 38317117]
22. Ajzen I. From intentions to actions: a theory of planned behavior. In: Kuhl J, Beckmann J, editors. Action Control. Berlin, Heidelberg: Springer; 1985.
23. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
24. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
25. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
26. Venkatesh V, Bala H. Technology Acceptance Model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
27. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117 [FREE Full text] [doi: 10.2196/13117] [Medline: 31115340]
28. Heinsch M, Wyllie J, Carlson J, Wells H, Tickner C, Kay-Lambkin F. Theories informing eHealth implementation: systematic review and typology classification. J Med Internet Res 2021 May 31;23(5):e18500 [FREE Full text] [doi: 10.2196/18500] [Medline: 34057427]
29. Rouidi M, Elouadi AE, Hamdoune A, Choujtani K, Chati A. TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: a systematic review. Inform Med Unlock 2022;32:101008. [doi: 10.1016/j.imu.2022.101008]
30. Yap YY, Tan SH, Choon SW. Elderly's intention to use technologies: a systematic literature review. Heliyon 2022 Jan;8(1):e08765 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e08765] [Medline: 35128090]
31. Kamal SA, Shafiq M, Kakria P. Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). Technol Soc 2020 Feb;60:101212. [doi: 10.1016/j.techsoc.2019.101212]
32. Yap YY, Tan SH, Tan SK, Choon SW. Integrating the capability approach and technology acceptance model to explain the elderly's use intention of online grocery shopping. Telematics Inform 2022 Aug;72:101842. [doi: 10.1016/j.tele.2022.101842]
33. Attié E, Meyer-Waarden L. The acceptance and usage of smart connected objects according to adoption stages: an enhanced technology acceptance model integrating the diffusion of innovation, uses and gratification and privacy calculus theories. Technol Forecast Soc Change 2022 Mar;176:121485. [doi: 10.1016/j.techfore.2022.121485]

34. Zhou M, Zhao L, Kong N, Campy KS, Qu S, Wang S. Factors influencing behavior intentions to telehealth by Chinese elderly: an extended TAM model. Int J Med Inform 2019 Jun;126:118-127. [doi: 10.1016/j.ijmedinf.2019.04.001] [Medline: 31029253]
35. Klingberg A, Sawe HR, Hammar U, Wallis LA, Hasselberg M. m-Health for burn injury consultations in a low-resource setting: an acceptability study among health care providers. Telemed J E Health 2020 Apr 01;26(4):395-405 [FREE Full text] [doi: 10.1089/tmj.2019.0048] [Medline: 31161967]
36. Tsai JM, Cheng MJ, Tsai HH, Hung SW, Chen YL. Acceptance and resistance of telehealth: the perspective of dual-factor concepts in technology adoption. Int J Inf Manag 2019 Dec;49:34-44. [doi: 10.1016/j.ijinfomgt.2019.03.003]
37. Steele R, Lo A, Secombe C, Wong YK. Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. Int J Med Inform 2009 Dec;78(12):788-801. [doi: 10.1016/j.ijmedinf.2009.08.001] [Medline: 19717335]
38. Katz E, Blumler JG, Gurevitch M. Utilization of mass communication by the individual. In: Blumler JG, Katz E, editors. The Uses of Mass Communications: Current Perspectives on Gratifications Research. Thousand Oaks, CA: Sage Publications; 1974:19-31.
39. Li C, Shi X, Dang J. Online communication and subjective well-being in Chinese college students: the mediating role of shyness and social self-efficacy. Comput Hum Behav 2014 May;34:89-95. [doi: 10.1016/j.chb.2014.01.032]
40. Yoon S. Does social capital affect SNS usage? A look at the roles of subjective well-being and social identity. Comput Hum Behav 2014 Dec;41:295-303. [doi: 10.1016/j.chb.2014.09.043]
41. Samuelson W, Zeckhauser R. Status quo bias in decision making. J Risk Uncertainty 1988 Mar;1(1):7-59. [doi: 10.1007/bf00055564]
42. Yap YY, Tan SH, Tan SK, Choon SW. Online grocery shopping intention: elderly's perspective in Malaysia. Heliyon 2023 Oct;9(10):e20827 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e20827] [Medline: 37916123]
43. Ajzen I, Fishbein M. A Bayesian analysis of attribution processes. Psychol Bull 1975 Mar;82(2):261-277. [doi: 10.1037/h0076477]
44. Park E, Kim KJ, Kwon SJ. Understanding the emergence of wearable devices as next-generation tools for health communication. Inf Technol People 2016 Nov 07;29(4):717-732. [doi: 10.1108/itp-04-2015-0096]
45. Papa A, Mital M, Pisano P, Del Giudice M. E-health and wellbeing monitoring using smart healthcare devices: an empirical investigation. Technol Forecast Soc Change 2020 Apr;153:119226 [FREE Full text] [doi: 10.1016/j.techfore.2018.02.018]
46. Rajak M, Shaw K. An extension of technology acceptance model for mHealth user adoption. Technol Soc 2021 Nov;67:101800. [doi: 10.1016/j.techsoc.2021.101800]
47. Ahn H, Park E. Determinants of consumer acceptance of mobile healthcare devices: An application of the concepts of technology acceptance and coolness. Telematics Inform 2022 May;70:101810. [doi: 10.1016/j.tele.2022.101810]
48. Yang HD, Yoo Y. It's all about attitude: revisiting the technology acceptance model. Decis Support Syst 2004 Oct;38(1):19-31. [doi: 10.1016/S0167-9236(03)00062-9]
49. Lazaro MJ, Lim J, Kim SH, Yun MH. Wearable technologies: acceptance model for smartwatch adoption among older adults. In: Proceedings of the 6th International Conference on Human Aspects of IT for the Aged Population. Technologies, Design and User Experience. 2020 Presented at: ITAP 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-50252-2_23]
50. Pan S, Jordan-Marsh M. Internet use intention and adoption among Chinese older adults: from the expanded technology acceptance model perspective. Comput Hum Behav 2010 Sep;26(5):1111-1119. [doi: 10.1016/j.chb.2010.03.015]
51. Braun MT. Obstacles to social networking website use among older adults. Comput Hum Behav 2013 May;29(3):673-680. [doi: 10.1016/j.chb.2012.12.004]
52. Li J, Ma Q, Chan AH, Man SS. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
53. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
54. Motamedi S, Masrahi A, Bopp T, Wang J. Different level automation technology acceptance: older adult driver opinion. Transp Res Part F 2021 Jul;80:1-13. [doi: 10.1016/j.trf.2021.03.010]
55. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
56. Kim HW, Kankanhalli A. Investigating user resistance to information systems implementation: a status quo bias perspective. MIS Q 2009;33(3):567-582. [doi: 10.2307/20650309]
57. Kim H. The effects of switching costs on user resistance to enterprise systems implementation. IEEE Trans Eng Manag 2011 Aug;58(3):471-482. [doi: 10.1109/tem.2010.2089630]
58. Hsieh PJ. An empirical investigation of patients' acceptance and resistance toward the health cloud: the dual factor perspective. Comput Hum Behav 2016 Oct;63:959-969. [doi: 10.1016/j.chb.2016.06.029]
59. Diener E. Subjective well-being. Psychol Bull 1984;95(3):542-575. [doi: 10.1037/0033-2909.95.3.542]
60. Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and longevity. Appl Psychol Health Well Being 2011;3(1):1-43. [doi: 10.1111/j.1758-0854.2010.01045.x]

61. Al-Jabri I, Sohail MS. Mobile banking adoption: application of diffusion of innovation theory. J Electron Commer Res 2012;13(4):379-391.
62. Wu J, Lu X. Effects of extrinsic and intrinsic motivators on using utilitarian, hedonic, and dual-purposed information systems: a meta-analysis. J Assoc Inf Syst 2013 Mar;14(3):153-191. [doi: 10.17705/1jais.00325]
63. Andreasen AR, Goldberg ME, Sirgy MJ. Foundational research on consumer welfare. In: Mick DG, Pettigrew S, Pechmann C, Ozanne JL, editors. Transformative Consumer Research for Personal and Collective Well-Being. New York, NY: Routledge; 2011.
64. Davis B, Pechmann C. Introduction to the special issue on transformative consumer research: developing theory to mobilize efforts that improve consumer and societal well-being. J Bus Res 2013 Aug;66(8):1168-1170. [doi: 10.1016/j.jbusres.2012.08.008]
65. Polites GL, Karahanna E. Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. MIS Q 2012;36(1):21-42. [doi: 10.2307/41410404]
66. Bem DJ. Self-perception theory. In: Gawronski B, editor. Advances in Experimental Social Psychology. Mahwah, NJ: Lawrence Erlbaum; 1972:1-62.
67. Petty RE, Cacioppo JT. Attitudes And Persuasion: Classic and Contemporary Approaches. Boulder, CO: Westview Press Inc; 1996.
68. Atchley RC. A continuity theory of normal aging. Gerontologist 1989 Apr 01;29(2):183-190. [doi: 10.1093/geront/29.2.183] [Medline: 2519525]
69. Venkatesh V. Determinants of perceived ease of use: integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
70. Zhou T. Examining location-based services usage from the perspectives of unified theory of acceptance and use of technology and privacy risk. J Electron Commer Res 2012;13(2):135-144.
71. Alaiad A, Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. Int J Med Inform 2014 Nov;83(11):825-840. [doi: 10.1016/j.ijmedinf.2014.07.003] [Medline: 25132284]
72. Wu IL, Li JY, Fu CY. The adoption of mobile healthcare by hospital's professionals: an integrative perspective. Decis Support Syst 2011 Jun;51(3):587-596. [doi: 10.1016/j.dss.2011.03.003]
73. Hong S, Tam K. Understanding the adoption of multipurpose information appliances: the case of mobile data services. Inf Syst Res 2006 Jun;17(2):162-179 [FREE Full text] [doi: 10.1287/isre.1060.0088]
74. Chang YZ, Ko CY, Hsiao CJ, Chen RJ, Yu CW, Cheng YW, et al. Understanding the determinants of implementing telehealth systems: a combined model of the theory of planned behavior and the technology acceptance model. J Appl Sci 2015 Jan 15;15(2):277-282. [doi: 10.3923/jas.2015.277.282]
75. Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS Q 2003;27(1):51-90. [doi: 10.2307/30036519]
76. Alam MZ, Hu W, Kaium MA, Hoque MR, Alam MM. Understanding the determinants of mHealth apps adoption in Bangladesh: a SEM-neural network approach. Technol Soc 2020 May;61:101255. [doi: 10.1016/j.techsoc.2020.101255]
77. Naik N, Ibrahim S, Sircar S, Patil V, Hameed BM, Rai BP, et al. Attitudes and perceptions of outpatients towards adoption of telemedicine in healthcare during COVID-19 pandemic. Ir J Med Sci 2022 Aug 16;191(4):1505-1512 [FREE Full text] [doi: 10.1007/s11845-021-02729-6] [Medline: 34402031]
78. Ateriya N, Saraf A, Meshram VP, Setia P. Telemedicine and virtual consultation: the Indian perspective. Natl Med J India 2018;31(4):215-218. [doi: 10.4103/0970-258x.258220]
79. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: a systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
80. Li X, Hess TJ, Valacich JS. Why do we trust new technology? A study of initial trust formation with organizational information systems. J Strateg Inf Syst 2008 Mar;17(1):39-71. [doi: 10.1016/j.jsis.2008.01.001]
81. Chew E, Teo SH, Tang WE, Ng DW, Koh GC, Teo VH. Trust and uncertainty in the implementation of a pilot remote blood pressure monitoring program in primary care: qualitative study of patient and health care professional views. JMIR Hum Factors 2023 Jan 05;10:e36072 [FREE Full text] [doi: 10.2196/36072] [Medline: 36602847]
82. Orrange S, Patel A, Mack WJ, Cassetta J. Patient satisfaction and trust in telemedicine during the COVID-19 pandemic: retrospective observational study. JMIR Hum Factors 2021 Apr 22;8(2):e28589 [FREE Full text] [doi: 10.2196/28589] [Medline: 33822736]
83. Zhang X, Zaman BU. Adoption mechanism of telemedicine in underdeveloped country. Health Informatics J 2020 Jun 30;26(2):1088-1103 [FREE Full text] [doi: 10.1177/1460458219868353] [Medline: 31566450]
84. Wu D, Gu H, Gu S, You H. Individual motivation and social influence: a study of telemedicine adoption in China based on social cognitive theory. Health Policy Technol 2021 Sep;10(3):100525. [doi: 10.1016/j.hlpt.2021.100525]
85. Ringle CM, Wende S, Becker JM. SmartPLS 3. SmartPLS GmbH. 2015. URL: https://www.smartpls.com [accessed 2023-10-12]

86. Chin WW, Marcolin BL, Newsted PR. A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Inf Syst Res 2003 Jun;14(2):189-217. [doi: 10.1287/isre.14.2.189.16018]
87. Wu ML. Structural Equation Model: Operation and Application of AMOS. Chongqing, China: Chongqing University Press; 2009.
88. Hair J, Hult GT, Ringle CM, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Thousand Oaks, CA: SAGE Publications; 2016.
89. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J Acad Mark Sci 2014 Aug 22;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
90. Franke G, Sarstedt M. Heuristics versus statistics in discriminant validity testing: a comparison of four procedures. Internet Res 2019 Jun 03;29(3):430-447. [doi: 10.1108/intr-12-2017-0515]
91. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur Bus Rev 2019 Jan 14;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
92. Ramayah T, Cheah JH, Chuah F, Ting H, Memon MA. Partial Least Squares Structural Equation Modeling (PLS-SEM) Using SmartPLS 3.0: An Updated and Practical Guide to Statistical Analysis. Kuala Lumpur, Malaysia: Pearson; 2016.
93. Hair J, Alamer A. Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: guidelines using an applied example. Res Methods Appl Linguist 2022 Dec;1(3):100027. [doi: 10.1016/j.rmal.2022.100027]
94. Hahn ED, Ang SH. From the editors: new directions in the reporting of statistical results in the Journal of World Business. J World Bus 2017 Feb;52(2):125-126. [doi: 10.1016/j.jwb.2016.12.003]
95. Shmueli G, Sarstedt M, Hair JF, Cheah JH, Ting H, Vaithilingam S, et al. Predictive model assessment in PLS-SEM: guidelines for using PLSpredict. Eur J Mark 2019 Nov 11;53(11):2322-2347. [doi: 10.1108/ejm-02-2019-0189]
96. Pires G, Lopes A, Correia P, Almeida L, Oliveira L, Panda R, et al. Usability of a telehealth solution based on TV interaction for the elderly: the VITASENIOR-MT case study. Univers Access Inf Soc 2023 Jan 17;22(2):525-536 [FREE Full text] [doi: 10.1007/s10209-021-00859-3] [Medline: 35069063]
97. Yang HL, Lin SL. The reasons why elderly mobile users adopt ubiquitous mobile social service. Comput Hum Behav 2019 Apr;93:62-75. [doi: 10.1016/j.chb.2018.12.005]
98. Ghani M, Adler C, Yeung H. Patient factors associated with interest in teledermatology: cross-sectional survey. JMIR Dermatol 2021 May 10;4(1):e21555 [FREE Full text] [doi: 10.2196/21555] [Medline: 37625162]
99. Goins RT, Kategile U, Dudley KC. Telemedicine, rural elderly, and policy issues. J Aging Soc Policy 2001 Oct 02;13(4):53-71. [doi: 10.1300/j031v13n04_04] [Medline: 12418382]

Abbreviations

HTMT: heterotrait-monotrait
IPMA: importance-performance map analysis
PLS-SEM: partial least squares structural equation modeling
TAM: technology acceptance model

Edited by Y Luo; submitted 26.05.24; peer-reviewed by F Abdelfattah, S Mohezar; comments to author 16.07.24; revised version received 24.10.24; accepted 29.11.24; published 24.03.25.

Please cite as:
Tan SH, Yap YY, Tan SK, Wong CK
Determinants of Telehealth Adoption Among Older Adults: Cross-Sectional Survey Study
JMIR Aging 2025;8:e60936
URL: https://aging.jmir.org/2025/1/e60936
doi:10.2196/60936
PMID:40126531

Using Machine Learning to Predict Cognitive Decline in Older Adults From the Chinese Longitudinal Healthy Longevity Survey: Model Development and Validation Study

Hao Ren[1,2*], PhD; Yiying Zheng[3*], MSc; Changjin Li[2*], MSc; Fengshi Jing[2], PhD; Qiting Wang[4], MSc; Zeyu Luo[2], BSc; Dongxiao Li[5], BSc; Deyi Liang[6], MSc; Weiming Tang[1,7,8], PhD; Li Liu[4], PhD; Weibin Cheng[1,2,9], PhD

[1]Institute for Healthcare Artificial Intelligence Application, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, No. 466 Xingangzhong Road, Haizhu District, Guangzhou, China

[2]Faculty of Data Science, City University of Macau, Macao SAR, China

[3]The Affiliated Traditional Chinese Medicine Hospital, Guangzhou Medical University, Guangzhou, China

[4]School of Public Health, Guangdong Pharmaceutical University, Guangzhou, China

[5]Hainan International College, Minzu University of China, Beijing, China

[6]Guangdong Women and Children Hospital, Guangzhou, China

[7]Institute for Global Health and Infectious Disease, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[8]University of North Carolina at Chapel Hill Project-China, Guangzhou, China

[9]College of Computing, City University of Hong Kong, Hong Kong SAR, China

[*]these authors contributed equally

Corresponding Author:
Weibin Cheng, PhD
Institute for Healthcare Artificial Intelligence Application, The Affiliated Guangdong Second Provincial General Hospital of Jinan University, No. 466 Xingangzhong Road, Haizhu District, Guangzhou, China

Abstract

Background: Cognitive impairment, indicative of Alzheimer disease and other forms of dementia, significantly deteriorates the quality of life of older adult populations and imposes considerable burdens on families and health care systems worldwide. The early identification of individuals at risk for cognitive impairment through a convenient and rapid method is crucial for the timely implementation of interventions.

Objective: The objective of this study was to explore the application of machine learning (ML) to integrate blood biomarkers, life behaviors, and disease history to predict the decline in cognitive function.

Methods: This approach uses data from the Chinese Longitudinal Healthy Longevity Survey. A total of 2688 participants aged 65 years or older from the 2008‐2009, 2011‐2012, and 2014 Chinese Longitudinal Healthy Longevity Survey waves were included, with cognitive impairment defined as a Mini-Mental State Examination (MMSE) score below 18. The dataset was divided into a training set (n=1331), an internal test set (n=333), and a prospective validation set (n=1024). Participants with a baseline MMSE score of less than 18 were excluded from the cohort to ensure a more accurate assessment of cognitive function. We developed ML models that integrate demographic information, health behaviors, disease history, and blood biomarkers to predict cognitive function at the 3-year follow-up point, specifically identifying individuals who are at risk of experiencing significant declines in cognitive function by that time. Specifically, the models aimed to identify individuals who would experience a significant decline in their MMSE scores (less than 18) by the end of the follow-up period. The performance of these models was evaluated using metrics including accuracy, sensitivity, and the area under the receiver operating characteristic curve.

Results: All ML models outperformed the MMSE alone. The balanced random forest achieved the highest accuracy (88.5% in the internal test set and 88.7% in the prospective validation set), albeit with a lower sensitivity, while logistic regression recorded the highest sensitivity. SHAP (Shapley Additive Explanations) analysis identified instrumental activities of daily living, age, and baseline MMSE scores as the most influential predictors for cognitive impairment.

Conclusions: The incorporation of blood biomarkers, along with demographic, life behavior, and disease history into ML models offers a convenient, rapid, and accurate approach for the early identification of older adult individuals at risk of cognitive impairment. This method presents a valuable tool for health care professionals to facilitate timely interventions and underscores the importance of integrating diverse data types in predictive health models.

(JMIR Aging 2025;8:e67437) doi:10.2196/67437

KEYWORDS

older adults; cognitive decline; Alzheimer disease; machine learning; blood biomarkers; disease history; Mini-Mental State Examination; MMSE; Chinese Longitudinal Healthy Longevity Survey; CLHLS

Introduction

Alzheimer disease (AD), the most prevalent form of dementia, is a progressive condition primarily characterized by memory loss [1]. It is estimated that approximately 50 million people worldwide are currently living with AD [2]. This condition not only deteriorates the quality of life for older adult individuals but also imposes significant burdens on families and health care systems, especially as the global population continues to age [3-6]. Mild cognitive impairment (MCI) is recognized as an intermediary stage between normal aging and the more severe cognitive decline observed in dementia. Early detection of MCI through diagnostic tools, such as magnetic resonance imaging (MRI), can facilitate timely interventions aimed at reducing the risk of progression to AD [7]. However, cognitive impairment, which encompasses a broader spectrum of decline, including MCI, is a critical concern that requires early identification. Predicting cognitive impairment, including MCI and other forms of decline, is essential for implementing preventive strategies and improving long-term health outcomes. Therefore, identifying individuals at potential risk for cognitive impairment is crucial.

While MRI and cerebrospinal fluid biomarkers, such as amyloid β, are significant indicators of AD [8], we aim to develop an algorithm that allows for the precise identification of individuals at risk using a more convenient and rapid approach, without relying on complex analyses like MRI or cerebrospinal fluid biomarkers. Recent studies have established a correlation between blood biomarkers and various factors, including diseases such as hypertension, diabetes, heart disease, and cerebrovascular disease, and lifestyle behaviors such as smoking, physical activity, and living conditions, with cognitive impairment [9,10]. High-dimensional data analysis has proven effective in capturing features that are crucial for identifying health issues [11,12]. Hence, the incorporation of these relevant factors into a sophisticated model allows for a practical application both in clinical settings and at home, enabling doctors or families to use existing data to train and monitor patients or older adult individuals, thereby identifying high-risk groups for further treatment and intervention.

In recent years, machine learning (ML), a branch of artificial intelligence, has been increasingly used for the prediction of disease outcome [13,14]. Unlike traditional methods that rely heavily on statistical significance, ML leverages algorithms to process existing factors and develop optimized models [15]. While there has been significant research into understanding the pathogenesis and influencing factors of MCI through ML, most of these studies have primarily focused on imaging techniques such as MRI [16]. Although MRI is a powerful diagnostic tool, its high cost and the inconvenience it poses limit its practicality for widespread use. In contrast, cognitive function assessments, such as the Mini-Mental State Examination (MMSE), offer a more accessible and feasible option for large-scale population screenings [17,18].

In this study, we used follow-up data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) collected during the years 2008‐2009, 2011‐2012 and 2014, encompassing a total of 2688 participants. Cognitive impairment was classified based on the MMSE scores, with a cutoff point set at 18, to determine cognitive status after 3 years. We included routine blood indices, lifestyle behaviors, and disease history from the baseline data in the ML model for training, aiming to predict the occurrence of cognitive impairment.

Methods

Study Participants

The cohort for this study is selected from CLHLS, which is a comprehensive longitudinal survey co-orchestrated by Peking University and the China Aging Science Research Center [19,20]. The survey targets Chinese seniors aged 65 years or older and includes detailed information about their living conditions, socioeconomic status, and health profiles [21]. The CLHLS initiates its baseline survey in 1998 and subsequently conducts follow-up surveys at regular intervals, with the relevant cohorts for this study being those from 2008‐2009, 2011‐2012, and 2014. The CLHLS study maintains ethical standards with approval from the Research Ethics Committee of Peking University (IRB00001052-13074), and all participants or their legal proxies provide written informed consent.

The initial samples from these specified years consist of 8418, 6066, and 3441 participants, respectively. Baseline subjects were screened by MMSE, and cognitive impairment was defined as an MMSE score <18 points [22]. The participants without missing data on MMSE scores at baseline and follow-up while with biomarker data were included in this study. This results in a narrowed-down research sample, with 602 participants from the 2008‐2009 cohort, 1263 from 2011‐2012, and 1116 from 2014, leaving a final sample size of 2688 subjects. Among the remaining participants, those from the 2008‐2009 and 2011‐2012 waves (n=1664) were further divided into a training set (n=1331) and an internal test set (n=333). The 1024 participants from the 2014 wave were used as a prospective validation set.

Predictors

Demographic predictors include age, gender, and BMI, calculated as weight in kilograms divided by height in meters squared (kg/m^2). Data on life behaviors and disease history are collected from the CLHLS questionnaire. Life behaviors account for living status (living alone or not), current smoking and drinking habits, exercise practices, marital status, and overall activity ability and mental health. Activity ability is assessed through activities of daily living (ADL) and instrumental activities of daily living (IADL). ADL was assessed by 6 indicators, including bathing, dressing, toileting, indoor transfer, continence, and eating. If all 6 items can be completely self-care, it means that daily life activities can take care of themselves. If

one or more items cannot be completely self-care, it means the daily life activities cannot take care of themselves completely (0=normal, 1=disability) [23,24]. IADL was assessed by 8 indicators, including visiting neighbors, going shopping, cooking a meal, washing clothing, walking continuously for 1 km at a time, lifting a weight of 5 kg, continuously crouching and standing up 3 times, and taking public transportation. If 8 items can be completed independently, such as the evaluation method of ADL, it means the instrumental daily life activities can be completed by themselves (0=normal or 1=disability) [23,25,26].

The development of the model in this study encompasses 3 categories of predictors derived from the baseline survey data: biomarkers, life behaviors, and disease history. The set of biomarkers comprises both routine blood examination indices and plasma biochemical examination indices. Routine blood indices include white blood cell count, red blood cell count, hemoglobin, erythrocyte hematocrit, erythrocyte mean corpuscular volume, erythrocyte mean corpuscular hemoglobin, erythrocyte mean corpuscular hemoglobin concentration, platelet count, plateletcrit, mean platelet volume, lymphocyte count, percentage of lymphocytes, and platelet distribution width. The plasma biochemical indices include high-density lipoprotein cholesterol (HDL), uric acid, plasma creatinine, glucose, triglyceride, total cholesterol (CHO), high-sensitivity c-reactive protein, malondialdehyde (MDA), and superoxide dismutase activity.

Mental health evaluation incorporates 7 questions, with 4 positively framed inquiries (regarding optimism, neatness, decision-making, and happiness relative to youth) and 3 negatively framed ones (concerning fear, loneliness, and feelings of decreased self-worth with age). Responses are scored on a scale, with higher scores correlating with poorer mental health. Disease history captures the presence or absence of hypertension, diabetes, heart disease, stroke, cancer, and arthritis. These multifaceted predictors collectively contribute to the ML model, providing a comprehensive profile for the assessment of cognitive impairment risk.

Data Preprocessing and Model Configuration

The data preprocessing phase involved addressing missing values and mitigating class imbalance to ensure a robust foundation for model training. Missing values in the dataset were imputed using the mean of the respective columns, which ensured completeness and preserved the statistical properties of the data [27]. In order to address the issue of class imbalance, the SMOTE (Synthetic Minority Over-Sampling Technique) was applied to the training set. SMOTE effectively generated synthetic samples for the minority class, enhancing the model's ability to learn from the imbalanced data distribution [28].

The selection of ML models was driven by a desire to compare different algorithms' performance and assess their robustness in dealing with imbalanced datasets. We chose 5 widely used algorithms: random forest (RF) [29], Extreme Gradient Boosting (XGBoost) [30], logistic regression [31], support vector machines (SVM) [32], and balanced random forest (BRF) [33]. These models were selected for their varied approaches to classification and their effectiveness in handling different types of data. RF and XGBoost are ensemble models that excel in handling high-dimensional data and capturing nonlinear relationships. Logistic regression and SVM are classic algorithms for binary classification, with SVM known for its ability to handle high-dimensional spaces effectively. The addition of BRF was made specifically to address the class imbalance issue in the RF model. The main difference between RF and BRF is how they handle class imbalance.

Each model was configured as follows:

- XGBoost (XGBClassifier): a learning rate of 0.03 and a maximum depth of 4 were chosen for the XGBoost model, with class weights incorporated to address class imbalance.
- Logistic regression: configured with a maximum of 1000 iterations and class weights set to "balanced" to account for class imbalance.
- SVM: probability estimation was enabled and class weights were balanced to ensure fairness across experiments.
- BRF: class weights were set to "balanced" to address the class imbalance within the dataset.

Outcomes

The assessment of cognitive function among participants was conducted using the MMSE, administered at both baseline and during follow-up sessions. The MMSE encompasses evaluations across various cognitive domains including orientation, registration, attention and calculation, recall, and language abilities, with a maximum achievable score of 30 points. In this study, the primary outcome for the ML model is the determination of cognitive impairment, defined as an MMSE score of less than 18 points at follow-up.

Model Construction, Evaluation, and Interpretation

The construction of each model followed a systematic approach to ensure rigorous evaluation and validation. Initially, the CLHLS 2008 - 2009 and CLHLS 2011 - 2012 datasets were merged. From this combined dataset, 1331/1664 (80%) of the data was randomly allocated as the training set, while the remaining 333/1664 (20%) was reserved as an internal test set for model evaluation. And then, the CLHLS 2014 dataset was used as a prospective validation set, with the model training based exclusively on the merged data from CLHLS 2008 - 2009 and CLHLS 2011 - 2012. Standard performance metrics—including accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve—were calculated to assess model effectiveness. All statistical analyses are conducted using Python (version 3.8; Python Software Foundation)

SHAP (Shapley Additive Explanations) [34] is a method for interpreting the output of ML models by assigning a contribution value to each feature, allowing us to understand the impact of individual predictors on a model's decision. In our study, SHAP was used to explain the predictive model for cognitive impairment in older adult individuals. SHAP values decompose the model's prediction into individual contributions from each feature, making it possible to attribute the output to the various risk factors in a transparent and interpretable manner.

Ethical Considerations

The CLHLS was approved by the Duke University Institutional Review Board (Pro00062871) and the Peking University Biomedical Ethics Committee (IRB00001052–13074). All participants provided written informed consent. The data used in this study were deidentified to protect participant privacy and confidentiality. No compensation wasprovided to participants.

Results

Participant Characteristics

Table 1 summarizes the baseline characteristics of the 2688 participants included in this study, stratified into 3 subsets: the training set (n=1331), the internal test set (n=333), and the prospective validation set (n=1024). Overall, the mean age was 79.73 (SD 10.97) years and slightly over half of the participants were male (53.24%). The mean BMI was 21.81 (SD 3.8 kg/m^2). Most participants lived with others (79.49%), did not smoke (77.79%), or drink (80.96%) at present, and had normal ADL (95.02%). Additionally, more than half (55.28%) showed abnormal mental health status, while 51.45% were married and living with a spouse. Several biochemical indicators were measured. The mean white blood cell count was 5.85 (SD 1.74×10^9)/L, red blood cell count was 4.64 (SD 1.68×10^{12})/L, and hemoglobin was 131.56 (SD 30.20) g/L. Participants had a mean HDL level of 1.31 (SD 0.38) mmol/L, and other biomarkers (eg, uric acid, creatinine, glucose, triglycerides, CHO, high-sensitivity c-reactive protein, MDA, and superoxide dismutase) were also assessed. Regarding disease history, hypertension was most prevalent (27.18%), followed by arthritis (11.31%), heart disease (8.01%), stroke (6.29%), and diabetes (2.75%). Cancer was reported by 0.42% of participants. Detailed distributions of these characteristics across the training, internal test, and prospective validation sets are presented in Table 1. A detailed flow diagram of participant inclusion and exclusion is provided in Figure 1.

Table . Characteristics of study subjects at baseline.

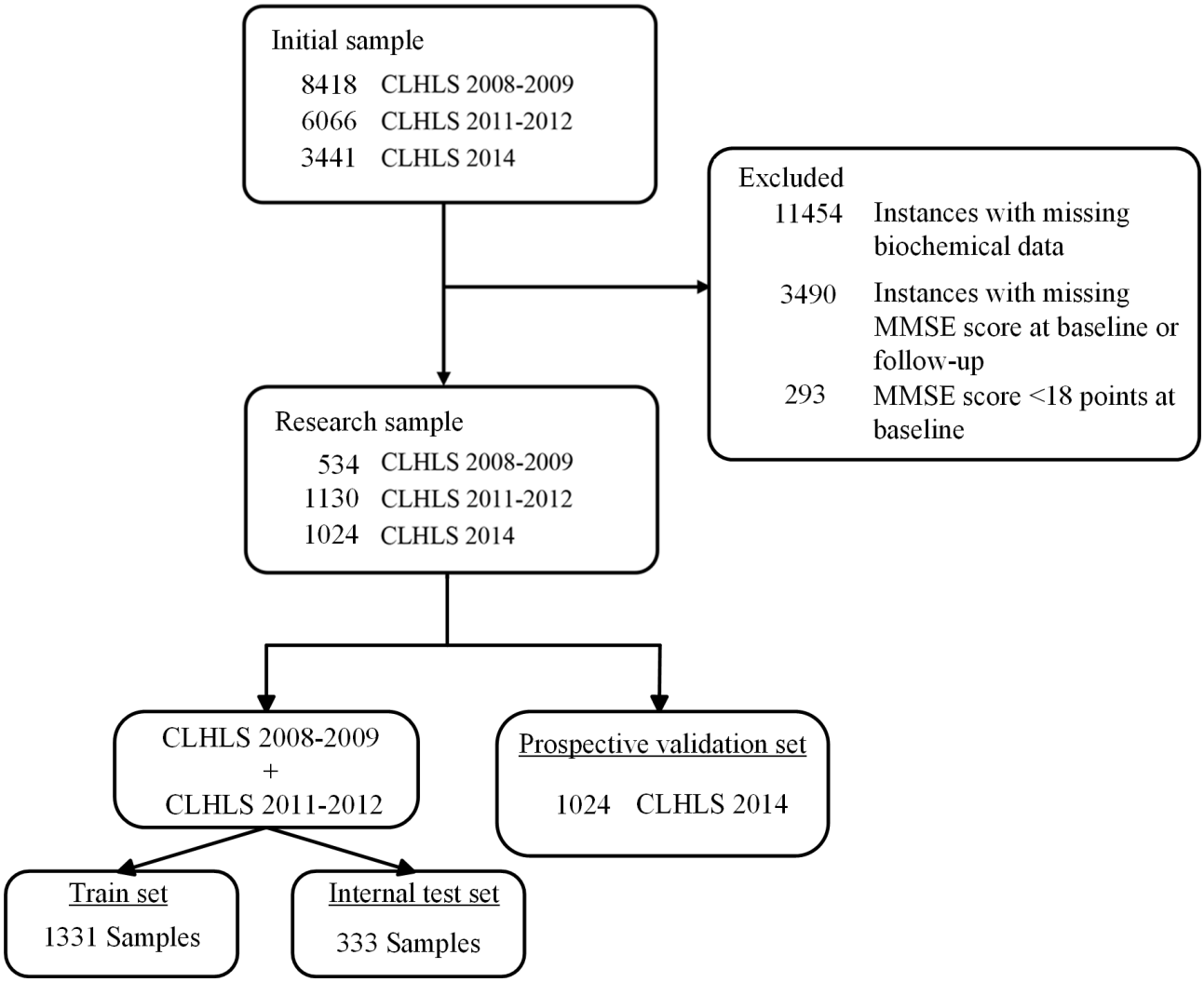
Predictors		Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
Demographics and life behaviors					
Age, year, mean (SD)		79.73 (10.97)	80.24 (11.43)	79.69 (11.6)	79.73 (10.07)
Gender, n (%)					
	Male	1431 (53.24)	685 (51.47)	173 (51.95)	573 (55.96)
	Female	1257 (46.76)	646 (48.53)	160 (48.05)	451 (44.04)
BMI, kg/m^2, mean (SD)		21.81 (3.8)	21.51 (3.86)	21.32 (3.67)	22.31 (3.64)
Living alone, n (%)					
	Yes	543 (20.51)	267 (20.37)	62 (18.9)	214 (21.21)
	No	2105 (79.49)	1044 (79.63)	266 (81.1)	795 (78.79)
Smoke at present, n (%)					
	Yes	593 (22.21)	317 (23.94)	72 (21.82)	204 (20.08)
	No	2077 (77.79)	1007 (76.06)	258 (78.18)	812 (79.92)
Drink at present, n (%)					
	Yes	509 (19.04)	265 (19.98)	57 (17.17)	187 (18.42)
	No	2164 (80.96)	1061 (80.02)	275 (82.83)	828 (81.58)
Exercise at present, n (%)					
	Yes	2134 (81.08)	1056 (81.23)	267 (81.65)	811 (80.7)
	No	498 (18.92)	244 (18.77)	60 (18.35)	194 (19.3)
ADL[a], n (%)					
	Normal	2479 (95.02)	1232 (93.97)	313 (95.72)	934 (96.19)
	Disability	130 (4.98)	79 (6.03)	14 (4.28)	37 (3.81)
IADL[b], n (%)					
	Normal	1593 (59.53)	742 (55.87)	198 (59.64)	653 (64.27)
	Disability	1083 (40.47)	586 (44.13)	134 (40.36)	363 (35.73)
Marital status, n (%)					
	Married and living with spouse	1370 (51.45)	654 (49.32)	165 (49.55)	551 (54.88)
	Others	1293 (48.55)	672 (50.68)	168 (20.45)	453 (45.12)
Mental health, n (%)					
	Normal	1122 (44.72)	542 (43.78)	139 (44.84)	441 (45.89)
	Abnormal	1387 (55.28)	696 (56.22)	171 (55.16)	520 (54.11)
Biomarkers, mean (SD)					
WBC[c], 10^9/L		5.85 (1.74)	5.63 (1.76)	5.64 (1.89)	6.19 (1.61)
RBC[d], 10^{12}/L		4.64 (1.68)	4.82 (2.11)	4.98 (1.9)	4.3 (0.56)
HGB[e], g/L		131.56 (30.2)	129.98 (26.56)	132.23 (22.69)	133.36 (36)
HCT[f], %		36.31 (15.05)	33.77 (18.01)	34.25 (18.1)	40.07 (7.23)
MCV[g], fL		93.95 (10.24)	93.51 (12.24)	92.94 (9.76)	94.81 (7.03)
MCH[h], pg		29.46 (7.24)	28.38 (5.25)	28.16 (5.75)	31.24 (9.22)
MCHC[i], g/L		312.18 (42.25)	304.34 (48.89)	301.38 (54.41)	325.58 (18.74)
PLT[j], 10^9/L		207.32 (99.24)	216.44 (109.67)	213.75 (111.53)	193.65 (77.22)

Predictors			Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
	PCT[k], %		0.27 (3.29)	0.33 (4.42)	0.2 (0.1)	0.18 (0.06)
	MPV[l], fL		9.54 (5.14)	9.43 (2.87)	9.97 (6.53)	9.55 (6.65)
	LYMPH[m], 10^9/L		14.56 (15.81)	23.13 (15.58)	21.28 (15.48)	2.03 (1.25)
	LYM%[n], %		19.55 (16.81)	10.49 (14.48)	11.59 (15.01)	33.06 (9.32)
	PDW[o], fL		15.47 (4.37)	14.86 (2.38)	14.87 (1.99)	16.44 (6.27)
	HDL[p], mmol/L		1.31 (0.38)	1.26 (0.37)	1.24 (0.38)	1.41 (0.37)
	UA[q], umol/L		291 (85.39)	285.55 (85.47)	288.51 (87.31)	298.83 (84.13)
	CRE[r], mmol/L		81.36 (24.19)	82.49 (25.97)	81.68 (22.2)	79.81 (22.28)
	GLU[s], mmol/L		5.06 (2.05)	4.83 (2.29)	4.85 (1.87)	5.42 (1.71)
	TG[u], mmol/L		1.29 (0.95)	1.22 (0.95)	1.44 (1.24)	1.34 (0.82)
	CHO[u], mmol/L		4.34 (1.17)	4.06 (1.18)	3.94 (1.22)	4.84 (0.95)
	CRPHS[v], mg/L		4.17 (18.46)	5.46 (23.21)	6.15 (29.35)	2.49 (5.29)
	MDA[w], nmol/ml		5.64 (3.4)	5.43 (2.76)	5.2 (2.63)	5.93 (4.04)
	SOD[x], IU/mL		55.53 (10.33)	55.53 (11.9)	53.34 (11.61)	56.04 (8.13)
Disease history, n (%)						
	Hypertension					
		Yes	716 (27.18)	316 (24.09)	88 (26.91)	312 (31.36)
		No	1918 (72.82)	996 (75.91)	239 (73.09)	683 (68.64)
	Diabetes					
		Yes	73 (2.75)	31 (2.34)	11 (3.33)	31 (3.11)
		No	2577 (97.25)	1292 (97.66)	319 (96.67)	966 (96.89)
	Heart disease					
		Yes	212 (8.01)	91 (6.9)	28 (8.54)	93 (9.29)
		No	2436 (91.99)	1228 (93.1)	300 (91.46)	908 (90.71)
	Stroke					
		Yes	167 (6.29)	77 (5.82)	20 (6.08)	70 (6.97)
		No	2490 (93.71)	1247 (94.18)	309 (93.92)	934 (93.03)
	Cancer					
		Yes	11 (0.42)	3 (0.23)	1 (0.3)	7 (0.73)
		No	2588 (99.58)	1301 (99.77)	329 (99.7)	958 (99.27)
	Arthritis					
		Yes	300 (11.31)	183 (13.86)	55 (16.72)	62 (6.18)

Predictors		Overall (n=2688)	Train set (n=1331)	Internal test set (n=333)	Prospective validation set (n=1024)
	No	2352 (88.69)	1137 (86.14)	274 (83.28)	941 (93.82)

[a]ADL: activities of daily living.
[b]IADL: instrumental activities of daily living.
[c]WBC: white blood cell count.
[d]RBC: red blood cell count.
[e]HGB: hemoglobin.
[f]HCT: erythrocyte hematocrit.
[g]MCV: erythrocyte mean corpuscular volume.
[h]MCH: erythrocyte mean corpuscular hemoglobin.
[i]MCHC: erythrocyte mean corpuscular hemoglobin concentration.
[j]PLT: platelet count.
[k]PCT: plateletocrit.
[l]MPV: mean platelet volume.
[m]LYMPH: lymphocyte count.
[n]LYM%: percentage of lymphocytes.
[o]PDW: platelet distribution width.
[p]HDL: high-density lipoprotein cholesterol.
[q]UA: urea acid.
[r]CRE: plasma creatine.
[s]GLU: glucose.
[t]TG: triglyceride.
[u]CHO: total cholesterol.
[v]CRPHS: high-sensitivity c-reactive protein.
[w]MDA: malondialdehyde.
[x]SOD: superoxide dismutase activity.

Figure 1. Flow diagram of the subject selection. CLHLS: Chinese Longitudinal Healthy Longevity Survey; MMSE: Mini-Mental State Examination.



Model Performance

The performance of the ML models in predicting follow-up cognitive impairment was evaluated using 5 algorithms—RF, XGBoost, logistic regression, SVM, and BRF—on both the internal test set and prospective validation set. Each model's accuracy, sensitivity, and area under the receiver operating characteristic curve were assessed. The detailed results are presented in Table 2 and visually depicted in Figure 2. Additionally, we used the MMSE as an input for ML prediction. After testing several ML models, the overall performance remained suboptimal. For illustration, we selected 1 representative result, as shown in Table 2.

Table . The prediction results for ML[a] models.

	Internal test			Prospective validation		
	Accuracy (95% CI)	Sensitivity (95% CI)	AUC[b] (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	AUC (95% CI)
RF[c]	0.828 (0.806-0.854)	0.688 (0.633-0.742)	0.81 (0.769-0.852)	0.813 (0.79-0.84)	0.684 (0.63-0.739)	0.806 (0.767-0.848)
XGBoost[d]	0.849 (0.825-0.87)	0.674 (0.621-0.728)	0.811 (0.769-0.851)	0.836 (0.812-0.857)	0.682 (0.629-0.733)	0.808 (0.768-0.848)
Logistic regression	0.778 (0.751-0.803)	0.715 (0.66-0.761)	0.803 (0.753-0.847)	0.771 (0.744-0.795)	0.745 (0.696-0.789)	0.804 (0.753-0.847)
SVM[e]	0.69 (0.661-0.719)	0.681 (0.629-0.731)	0.777 (0.725-0.827)	0.672 (0.643-0.701)	0.676 (0.623-0.725)	0.777 (0.726-0.826)
Balanced RF classifier	0.885 (0.866-0.903)	0.58 (0.54-0.626)	0.809 (0.765-0.849)	0.887 (0.867-0.905)	0.616 (0.572-0.664)	0.811 (0.767-0.849)
MMSE[f]	N/A[g]	N/A	0.571 (0.485-0.653)	N/A	N/A	0.558 (0.494-0.62)

[a]ML: machine learning.

[b]AUC: area under the curve.

[c]RF: random forest.

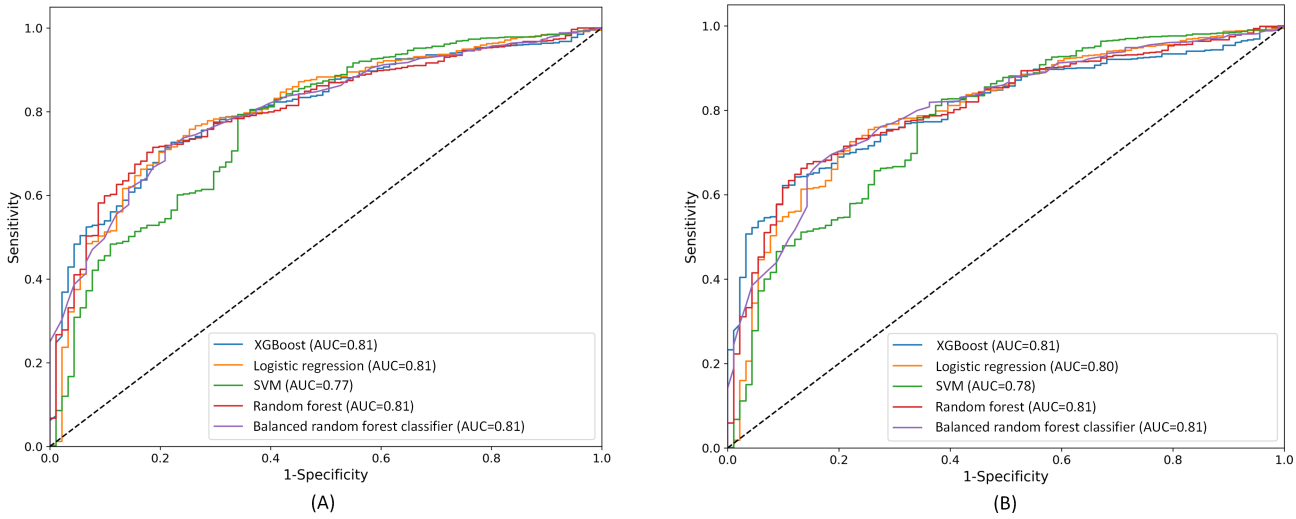
[d]XGBoost: Extreme Gradient Boosting.

[e]SVM: support vector machines.

[f]MMSE: Mini-Mental State Examination.

[g]N/A: not applicable.

Figure 2. ROC curves with AUC values for machine learning models: (A) internal test set and (B) prospective validation set. AUC: area under the curve; ROC: receiver operating characteristic; SVM: support vector machines; XGBoost: Extreme Gradient Boosting.



In the internal test set, the BRF achieved the highest accuracy at 88.5% but had lower sensitivity at 58%. RF and XGBoost provided balanced results with accuracies of 82.8% and 84.9% and sensitivities of 68.8% and 67.4%, respectively. Logistic regression yielded a moderate accuracy of 77.8% but the highest sensitivity at 71.5%, while SVM had the lowest performance with 69% accuracy and 68.1% sensitivity. Similar patterns were observed in the prospective validation set, with BRF at 88.7% accuracy, logistic regression reaching 74.5% sensitivity, and SVM again showing the lowest overall performance. Additionally, the performance of all ML models was superior to that of the MMSE.

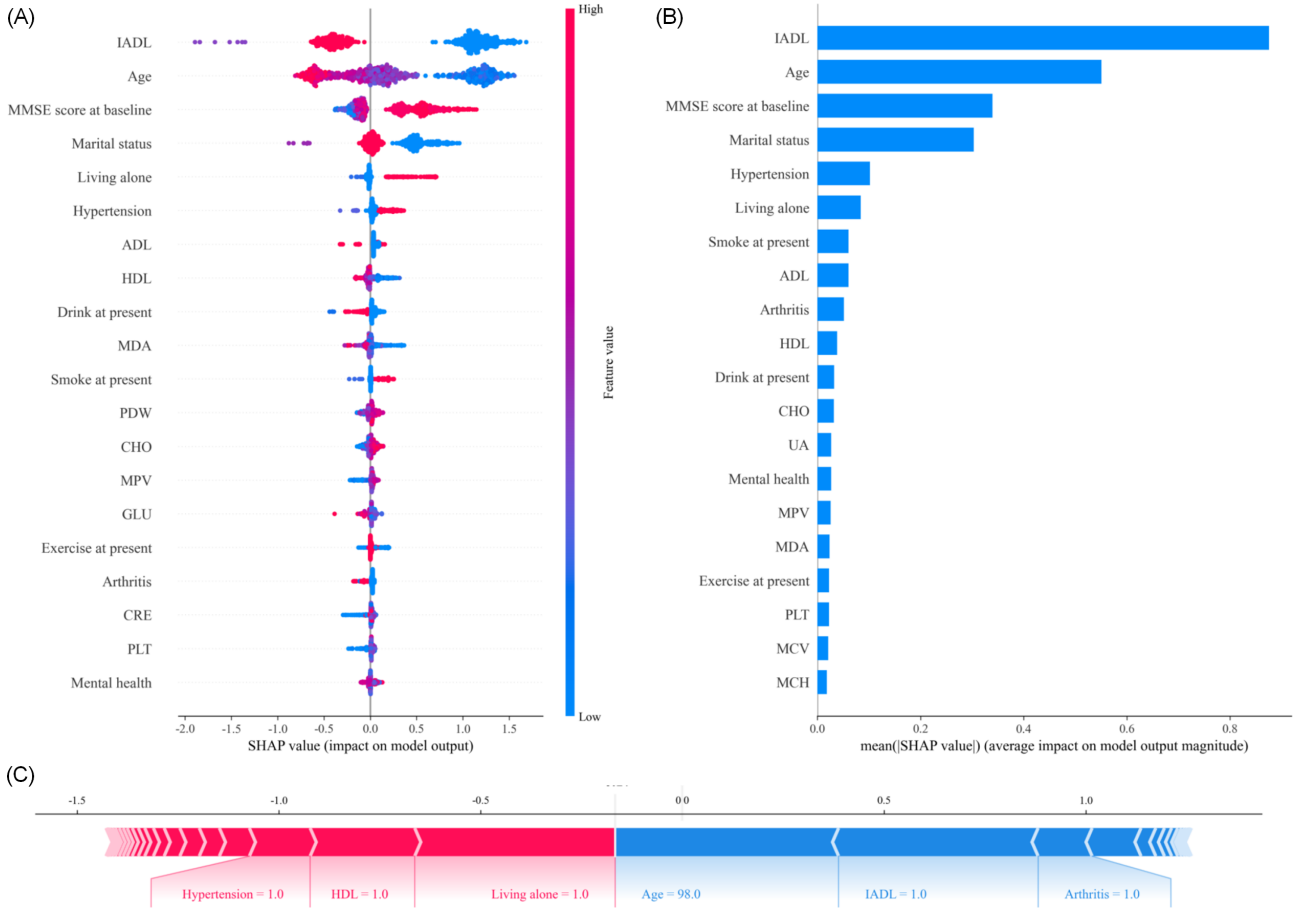
Figure 3 offers a detailed exposition of the XGBoost for cognitive impairment in older adult individuals, featuring the 20 most influential risk predictors. The model prioritizes IADL, age, and baseline MMSE scores as the top determinants, with marital status, living alone, and hypertension also providing significant predictive value. Other important factors include biological markers such as MDA and HDL, alongside lifestyle factors such as current exercise, smoking, and drinking habits.

Figure 3A illustrates the distribution of SHAP values for these predictors, indicating their impact on the XGBoost's output. In this plot, each dot represents a single instance of a feature in the dataset, and the horizontal axis shows the SHAP value of that feature. The SHAP value indicates the contribution of each feature to the model's prediction, with positive values suggesting that the feature increases the likelihood of cognitive impairment, and negative values indicating a decrease. The color of the dots, ranging from blue to red, represents the feature's value, with blue corresponding to lower values and red indicating higher values of the predictor. This color scheme helps highlight how different values of each predictor influence the model's outcome.

A high SHAP value for a given feature corresponds to a high level of importance in the predictive model. Figure 3B features a bar plot that quantifies the average impact of each predictor, measured by the mean absolute SHAP value. The length of each bar represents the average contribution of a feature to the model's output across all data points. Here, the most important features in the model—such as IADL, age, and MMSE score at baseline—are easily identified, as they have the longest bars, indicating that they have the highest average impact on the model's predictions.

Figure 3C illustrates a specific case study, showing how the SHAP values for a particular individual (in this case, an adult aged 98 years) contribute to the XGBoost's prediction of cognitive impairment. Each feature is shown with its value (eg, hypertension=1), and the arrows indicate how these values shift the model's prediction. Features such as hypertension and living alone appear to have a red color, indicating they push the prediction toward a higher risk of cognitive impairment. Similarly, age (with a value of 98) and IADL further emphasize the risk in this individual's profile. The interaction of these predictors is visualized through their SHAP values, which collectively guide the prediction model's decision, offering an individualized risk profile for this person.

Figure 3. Explanation of the interpretability of the XGBoost (Extreme Gradient Boosting) model (the best-performing model) for predicting older adult mortality. (A) and (B) show the top 20 risk predictors for prediction of cognitive impairment subjects, and (C) shows the SHAP plots of a subject. ADL: activities of daily living; CHO: total cholesterol; CRE: plasma creatine; GLU: glucose; HDL: high-density lipoprotein cholesterol; IADL: instrumental activities of daily living; MCH: erythrocyte mean corpuscular hemoglobin; MCV: erythrocyte mean corpuscular volume; MDA: malondialdehyde; MMSE: Mini-Mental State Examination; MPV: mean platelet volume; PDW: platelet distribution width; PLT: platelet count; SHAP: Shapley Additive Explanations; UA: uric acid.



Discussion

Principal Findings

This study uses the capabilities of ML to integrate a diverse set of 39 predictors for forecasting the decline in cognitive function over a 3-year period, yielding significant findings that resonate with the existing body of literature. Our ML model underscores the importance of IADL, age, baseline MMSE scores, marital status, living arrangements, hypertension, arthritis, and general lifestyle habits as pivotal factors influencing cognitive function. These determinants are consistent with findings from prior research, thereby affirming the reliability and relevance of our analytical approach [12,22].

Consistent with prior studies, advanced age has been identified as a significant risk factor for cognitive impairment [35,36]. The risk of MCI in the older adults aged 65 years or older is as high as 10% - 20% [37]. The limitation of older adult individuals' ability to perform activities, as assessed by ADL and IADL, restricts their range of activities, diminishes social interactions, and consequently reduces the cerebral stimulation necessary for maintaining cognitive functions [38,39]. Furthermore, the impact of living conditions on cognitive health

is evidenced by the predictive value of marital status and living alone [22,40-42]. Older adult individuals residing with a spouse typically exhibit healthier brain functions due to increased communication and maintenance of a normal life.

Additionally, this study underscores the detrimental effects of unhealthy lifestyle behaviors such as smoking and drinking on cognitive function [43-45], as well as the risk of cognitive impairment caused by pre-existing health conditions [46-49]. For instance, an Indian study highlighted that older smokers were 24% more likely to experience cognitive impairment compared to nonsmokers [50]. Similarly, Sabia et al [45] reported that abstaining from alcohol or consuming more than 14 units per week in middle age escalates the risk of AD. Moreover, recent research suggests that up to 3% of dementia cases could be averted by enhancing physical activity levels [51,52]. Concerning health conditions, a community-based cohort study illustrated that hypertension is associated with an increased risk of both all-cause MCI and nonamnestic MCI, with hazard ratios of 1.4 and 1.7, respectively, after age and sex adjustments [49]. Additionally, Appenzeller et al [53] found that patients with rheumatoid arthritis exhibited a significantly higher incidence of cognitive impairment compared to healthy controls.

In this study, in addition to age, lifestyle behaviors, and disease history—which can be evaluated through questionnaires or scales—biomarkers were specifically included in the ML model to enhance the predictive accuracy of clinical risk assessments for cognitive decline. Prior research has indicated that count elevated levels of specific biomarkers, including MDA, HDL, platelets, mean platelet volume, platelet distribution width, mean corpuscular hemoglobin, CHO, lymphocyte percentage, and plasma creatinine, are associated with an increased risk of cognitive decline [54-65]. For clinicians, integrating a patient's lifestyle behaviors with blood biochemical markers can aid in the assessment of cognitive function. For communities, these indicators can help identify residents who may be at high risk. For family members, this model enables the evaluation of older adult relatives who may be reluctant to acknowledge their cognitive decline, thereby facilitating timely medical intervention. The model developed in this study is versatile and offers valuable insights for the identification of cognitive impairment across various settings.

The ML model developed in this study has the potential to significantly improve clinical practice and primary care by providing a rapid, efficient, and accessible tool for identifying individuals at risk of cognitive decline. Traditional methods for cognitive function assessment, such as imaging techniques such as MRI, can be time-consuming and resource-intensive, especially in resource-limited settings. By leveraging routine data such as blood biomarkers, demographic information, and lifestyle factors, this model offers a cost-effective approach to identify individuals who may require further clinical evaluation or early intervention. In primary care settings, where health care professionals often manage large volumes of patients, the model can serve as a valuable screening tool to detect early cognitive decline and facilitate referrals for specialized care. Furthermore, by integrating this model into electronic health records, health care providers can make timely and informed decisions, improving patient outcomes through proactive management. In essence, the model has the potential to transform early detection and intervention strategies, shifting the focus toward preventative care and better allocation of health care resources.

Compared to previous studies using ML for cognitive impairment prediction, our study offers several distinct contributions. For example, while studies such as those by Hu et al [22] and Gao et al [66] have successfully developed ML models to predict cognitive impairment among Chinese community-dwelling older adult individuals, they often focused on a more limited set of predictors—typically emphasizing demographic factors and neuropsychological assessments. In contrast, our study integrates a comprehensive set of 39 predictors, including both routine blood biomarkers (eg, MDA and HDL) and detailed lifestyle and disease history data. This broader approach not only enhances predictive accuracy but also provides a rapid, cost-effective tool that can be easily applied in community and clinical settings. Moreover, while some previous work [67,68] has primarily relied on imaging data or traditional statistical methods, our use of advanced ensemble ML techniques (such as BRF and XGBoost) combined with SHAP-based interpretability offers a clearer understanding of individual risk factors. This interpretability is crucial for clinicians to tailor early intervention strategies. In summary, our study advances the field by delivering a more inclusive and interpretable model that effectively tracks cognitive decline over 3 years, thereby addressing gaps in existing research and offering tangible benefits for early detection and intervention.

Nevertheless, this study presents certain limitations that warrant consideration. Primarily, the dependence on MMSE scores as the sole measure of cognitive impairment may not fully represent the broad spectrum of cognitive health, as the MMSE mainly evaluates specific cognitive domains and does not address emotional or psychological aspects. Furthermore, there is potential bias due to the generally younger age and better overall function of this study's participants compared to those who were lost to follow-up, which could skew the results. However, the representation of the dataset at a national level does provide a measure of balance, helping to partially offset these biases.

Conclusions

In conclusion, this study validates the efficacy of a ML model integrating demographic data, lifestyle factors, and biomarkers to predict cognitive impairment in older adults. It underscores the significance of traditional risk factors such as age and daily functional abilities while highlighting the role of solitary living conditions and unhealthy habits in cognitive decline. By including a broad spectrum of biomarkers, the model enriches the predictive framework, offering clinicians, communities, and families a valuable tool for early identification and intervention in cognitive impairment, which could have far-reaching implications for public health and the well-being of the aging population.

Limitations

While the CLHLS is a large-scale longitudinal study that primarily focuses on individuals aged 65 years and older, to assess the health status and longevity of the older adult

population in China, it is important to acknowledge the limitations of the dataset concerning its representativeness. The CLHLS sample is designed to represent the health conditions of the older adult population in China, but it may not fully capture the global demographics of AD or other forms of dementia. Specifically, older adult populations in other countries or regions may differ by genetic background, lifestyle factors, and health risks, which could influence the development and progression of cognitive impairment. Future studies incorporating diverse, multinational cohorts would be beneficial in enhancing the generalizability and robustness of cognitive decline prediction models.

Acknowledgments

We thank CLHLS for providing data, and all participants for supporting this study. This study was supported by Guangzhou Science and Technology Plan Project (2024A03J1074 and 2024A03J1067) and General Project of Guangdong Provincial Natural Science Foundation (2025A1515012013).

Data Availability

The data used in this study are stored at Peking University and available upon request.

Authors' Contributions

YZ and CL contributed equally to the work. YZ, CL, HR, QW, ZL, D Li, D Liang, WT, and LL performed the data collection and analysis. YZ, CL, HR, and QW contributed to visualization. YZ and HR were responsible for data validation. YZ, CL and HR wrote this paper, and review editing was performed by HR and WC. All the authors read and approved the final paper.

Conflicts of Interest

None declared.

References

1. Hane FT, Robinson M, Lee BY, Bai O, Leonenko Z, Albert MS. Recent progress in Alzheimer's disease research, part 3: diagnosis and treatment. J Alzheimers Dis 2017;57(3):645-665. [doi: 10.3233/JAD-160907] [Medline: 28269772]
2. Scheltens P, De Strooper B, Kivipelto M, et al. Alzheimer's disease. Lancet 2021 Apr 24;397(10284):1577-1590. [doi: 10.1016/S0140-6736(20)32205-4] [Medline: 33667416]
3. 2023 Alzheimer's disease facts and figures. Alzheimers Dement 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]
4. Prince M, Wimo A, Ali GC, Wu YT, Prina M. World Alzheimer report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. : Alzheimer's Disease International; 2015.
5. Takizawa C, Thompson PL, van Walsem A, Faure C, Maier WC. Epidemiological and economic burden of Alzheimer's disease: a systematic literature review of data across Europe and the United States of America. JAD 2014;43(4):1271-1284. [doi: 10.3233/JAD-141134]
6. Jia J, Wei C, Chen S, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement 2018 Apr;14(4):483-491. [doi: 10.1016/j.jalz.2017.12.006] [Medline: 29433981]
7. Anderson ND. State of the science on mild cognitive impairment (MCI). CNS Spectr 2019 Feb;24(1):78-87. [doi: 10.1017/S1092852918001347] [Medline: 30651152]
8. Tan CC, Yu JT, Tan L. Biomarkers for preclinical Alzheimer's disease. J Alzheimers Dis 2014;42(4):1051-1069. [doi: 10.3233/JAD-140843] [Medline: 25024325]
9. Dong X, Nao J, Shi J, Zheng D. Predictive value of routine peripheral blood biomarkers in Alzheimer's disease. Front Aging Neurosci 2019;11:332. [doi: 10.3389/fnagi.2019.00332] [Medline: 31866854]
10. Jia L, Du Y, Chu L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: a cross-sectional study. Lancet Public Health 2020 Dec;5(12):e661-e671. [doi: 10.1016/S2468-2667(20)30185-7] [Medline: 33271079]
11. Jing F, Cheng M, Li J, et al. Social, lifestyle, and health status characteristics as a proxy for occupational burnout identification: a network approach analysis. Front Psychiatry 2023;14:1119421. [doi: 10.3389/fpsyt.2023.1119421] [Medline: 37124263]
12. Tan WY, Hargreaves C, Chen C, Hilal S. A machine learning approach for early diagnosis of cognitive impairment using population-based data. J Alzheimers Dis 2023;91(1):449-461. [doi: 10.3233/JAD-220776] [Medline: 36442196]
13. Aguayo GA, Zhang L, Vaillant M, et al. Machine learning for predicting neurodegenerative diseases in the general older population: a cohort study. BMC Med Res Methodol 2023 Jan 11;23(1):8. [doi: 10.1186/s12874-023-01837-4] [Medline: 36631766]
14. Tan M, Xiao Y, Jing F, et al. Evaluating machine learning-enabled and multimodal data-driven exercise prescriptions for mental health: a randomized controlled trial protocol. Front Psychiatry 2024;15:1352420. [doi: 10.3389/fpsyt.2024.1352420] [Medline: 38287940]

15. Navarro CLA, Damen JAA, Takada T, et al. Completeness of reporting of clinical prediction models developed using supervised machine learning: a systematic review. BMC Med Res Methodol 2022 Jan 13;22(1):12. [doi: 10.1186/s12874-021-01469-6] [Medline: 35026997]
16. Frizzell TO, Glashutter M, Liu CC, et al. Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: a systematic review. Ageing Res Rev 2022 May;77:101614. [doi: 10.1016/j.arr.2022.101614] [Medline: 35358720]
17. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: a comprehensive review. J Am Geriatr Soc 1992 Sep;40(9):922-935. [doi: 10.1111/j.1532-5415.1992.tb01992.x] [Medline: 1512391]
18. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6]
19. The Chinese Longitudinal Healthy Longevity Survey (CLHLS)—Longitudinal Data (1998–2018). Peking University Open Research Data Platform. 1998. URL: https://opendata.pku.edu.cn/dataset.xhtml?persistentId=doi:10.18170/DVN/WBO7LK [accessed 2025-04-19]
20. Zeng Y, Vaupel JW. Chinese Longitudinal Healthy Longevity Survey (CLHLS), biomarkers datasets, 2009, 2012, 2014 (ICPSR 37226). National Archive of Computerized Data on Aging. 2009. URL: https://www.icpsr.umich.edu/web/NACDA/studies/37226/versions/V1 [accessed 2025-04-19]
21. Wu J, Li J. The impact of social participation on older people's death risk: an analysis from CLHLS. China Popul Dev Stud 2018 Jun;2(2):173-185. [doi: 10.1007/s42379-018-0008-x]
22. Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among chinese community-dwelling elderly people with normal cognition: development and validation study. J Med Internet Res 2021 Feb 24;23(2):e20298. [doi: 10.2196/20298] [Medline: 33625369]
23. Mlinac ME, Feng MC. Assessment of activities of daily living, self-care, and independence. Arch Clin Neuropsychol 2016 Sep;31(6):506-516. [doi: 10.1093/arclin/acw049] [Medline: 27475282]
24. Farias ST, Mungas D, Reed BR, Harvey D, Cahn-Weiner D, Decarli C. MCI is associated with deficits in everyday functioning. Alzheimer Dis Assoc Disord 2006;20(4):217-223. [doi: 10.1097/01.wad.0000213849.51495.d9] [Medline: 17132965]
25. Kiosses DN, Alexopoulos GS. IADL functions, cognitive deficits, and severity of depression: a preliminary study. Am J Geriatr Psychiatry 2005 Mar;13(3):244-249. [doi: 10.1176/appi.ajgp.13.3.244] [Medline: 15728756]
26. Graf C. The Lawton instrumental activities of daily living scale. AJN 2008;108(4):52-62. [doi: 10.1097/01.NAJ.0000314810.46029.74]
27. Little RJA, Rubin D. Statistical Analysis with Missing Data: John Wiley & Sons; 2019.
28. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. JAIR 2002;16:321-357. [doi: 10.1613/jair.953]
29. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
30. Chen T, Guestrin C. XGBoost. Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13-17, 2016; San Francisco, CA, United States.
31. Sturdivant RX, Hosmer DW, Lemeshow S. Applied Logistic Regression: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
32. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
33. Khalilia M, Chakraborty S, Popescu M. Predicting disease risks from highly imbalanced data using random forest. BMC Med Inform Decis Mak 2011 Jul 29;11:1-13. [doi: 10.1186/1472-6947-11-51] [Medline: 21801360]
34. Mangalathu S, Hwang SH, Jeon JS. Failure mode and effects analysis of RC members based on machine-learning-based Shapley Additive Explanations (SHAP) approach. Eng Struct 2020 Sep;219:110927. [doi: 10.1016/j.engstruct.2020.110927]
35. Cong L, Ren Y, Wang Y, et al. Mild cognitive impairment among rural-dwelling older adults in China: a community-based study. Alzheimers Dement 2023 Jan;19(1):56-66. [doi: 10.1002/alz.12629] [Medline: 35262288]
36. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimers Res Ther 2021 Sep 28;13(1):162. [doi: 10.1186/s13195-021-00900-w] [Medline: 34583745]
37. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA 2014 Dec 17;312(23):2551-2561. [doi: 10.1001/jama.2014.13806] [Medline: 25514304]
38. Njegovan V, Hing MM, Mitchell SL, Molnar FJ. The hierarchy of functional loss associated with cognitive decline in older persons. J Gerontol A Biol Sci Med Sci 2001 Oct;56(10):M638-M643. [doi: 10.1093/gerona/56.10.m638] [Medline: 11584037]
39. Jekel K, Damian M, Wattmo C, et al. Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review. Alzheimers Res Ther 2015;7(1):17. [doi: 10.1186/s13195-015-0099-0] [Medline: 25815063]
40. Sundström A, Westerlund O, Kotyrlo E. Marital status and risk of dementia: a nationwide population-based prospective study from Sweden. BMJ Open 2016 Jan 4;6(1):e008565. [doi: 10.1136/bmjopen-2015-008565] [Medline: 26729377]
41. Sommerlad A, Ruegger J, Singh-Manoux A, Lewis G, Livingston G. Marriage and risk of dementia: systematic review and meta-analysis of observational studies. J Neurol Neurosurg Psychiatry 2018 Mar;89(3):231-238. [doi: 10.1136/jnnp-2017-316274] [Medline: 29183957]

42. Ahn IS, Kim JH, Kim S, et al. Impairment of instrumental activities of daily living in patients with mild cognitive impairment. Psychiatry Investig 2009 Sep;6(3):180-184. [doi: 10.4306/pi.2009.6.3.180] [Medline: 20046393]
43. Tyas SL, White LR, Petrovitch H, et al. Mid-life smoking and late-life dementia: the Honolulu-Asia aging study. Neurobiol Aging 2003;24(4):589-596. [doi: 10.1016/s0197-4580(02)00156-2] [Medline: 12714116]
44. Xu W, Wang H, Wan Y, et al. Alcohol consumption and dementia risk: a dose-response meta-analysis of prospective studies. Eur J Epidemiol 2017 Jan;32(1):31-42. [doi: 10.1007/s10654-017-0225-3] [Medline: 28097521]
45. Sabia S, Fayosse A, Dumurgier J, et al. Alcohol consumption and risk of dementia: 23 year follow-up of Whitehall II cohort study. BMJ 2018 Aug 1;362:k2927. [doi: 10.1136/bmj.k2927] [Medline: 30068508]
46. Vitturi BK, Nascimento BAC, Alves BR, de Campos FSC, Torigoe DY. Cognitive impairment in patients with rheumatoid arthritis. J Clin Neurosci 2019 Nov;69:81-87. [doi: 10.1016/j.jocn.2019.08.027] [Medline: 31447371]
47. Meade T, Manolios N, Cumming SR, Conaghan PG, Katz P. Cognitive impairment in rheumatoid arthritis: a systematic review. Arthritis Care Res (Hoboken) 2018 Jan;70(1):39-52. [doi: 10.1002/acr.23243] [Medline: 28371512]
48. Aronow WS. Hypertension and cognitive impairment. Ann Transl Med 2017 Jun;5(12):259. [doi: 10.21037/atm.2017.03.99] [Medline: 28706927]
49. Reitz C, Tang MX, Manly J, Mayeux R, Luchsinger JA. Hypertension and the risk of mild cognitive impairment. Arch Neurol 2007 Dec;64(12):1734-1740. [doi: 10.1001/archneur.64.12.1734] [Medline: 18071036]
50. Muhammad T, Govindu M, Srivastava S. Relationship between chewing tobacco, smoking, consuming alcohol and cognitive impairment among older adults in India: a cross-sectional study. BMC Geriatr 2021 Jan 29;21(1):85. [doi: 10.1186/s12877-021-02027-x] [Medline: 33514331]
51. Liang JH, Lu L, Li JY, et al. Contributions of modifiable risk factors to dementia incidence: a Bayesian network analysis. J Am Med Dir Assoc 2020 Nov;21(11):1592-1599. [doi: 10.1016/j.jamda.2020.04.006] [Medline: 32563753]
52. Dominguez LJ, Veronese N, Vernuccio L, et al. Nutrition, physical activity, and other lifestyle factors in the prevention of cognitive decline and dementia. Nutrients 2021 Nov 15;13(11):4080. [doi: 10.3390/nu13114080] [Medline: 34836334]
53. Appenzeller S, Bertolo MB, Costallat LTL. Cognitive impairment in rheumatoid arthritis. Methods Find Exp Clin Pharmacol 2004 Jun;26(5):339-343. [doi: 10.1358/mf.2004.26.5.831324] [Medline: 15319812]
54. Keller JN, Schmitt FA, Scheff SW, et al. Evidence of increased oxidative damage in subjects with mild cognitive impairment. Neurology (ECronicon) 2005 Apr 12;64(7):1152-1156. [doi: 10.1212/01.WNL.0000156156.13641.BA] [Medline: 15824339]
55. Liu Z, Liu Y, Tu X, et al. High serum levels of malondialdehyde and 8-OHdG are both associated with early cognitive impairment in patients with acute ischaemic stroke. Sci Rep 2017 Aug 25;7(1):9493. [doi: 10.1038/s41598-017-09988-3] [Medline: 28842715]
56. van Exel E, de Craen AJM, Gussekloo J, et al. Association between high-density lipoprotein and cognitive impairment in the oldest old. Ann Neurol 2002 Jun;51(6):716-721. [doi: 10.1002/ana.10220] [Medline: 12112077]
57. Atzmon G, Gabriely I, Greiner W, Davidson D, Schechter C, Barzilai N. Plasma HDL levels highly correlate with cognitive function in exceptional longevity. J Gerontol A Biol Sci Med Sci 2002 Nov;57(11):M712-M715. [doi: 10.1093/gerona/57.11.m712] [Medline: 12403798]
58. Crichton GE, Elias MF, Davey A, Sullivan KJ, Robbins MA. Higher HDL cholesterol is associated with better cognitive function: the Maine-Syracuse study. J Int Neuropsychol Soc 2014 Nov;20(10):961-970. [doi: 10.1017/S1355617714000885] [Medline: 25382185]
59. Zhang Y, Liu J, Wei Z, et al. Elevated serum platelet count inhibits the effects of brain functional changes on cognitive function in patients with mild cognitive impairment: A resting-state functional magnetic resonance imaging study. Front Aging Neurosci 2023;15:1088095. [doi: 10.3389/fnagi.2023.1088095] [Medline: 37051376]
60. Sun D, Wang Q, Kang J, et al. Correlation between serum platelet count and cognitive function in patients with atrial fibrillation: a cross-sectional study. Cardiol Res Pract 2021;2021:9039610. [doi: 10.1155/2021/9039610] [Medline: 33824766]
61. Wang RT, Jin D, Li Y, Liang QC. Decreased mean platelet volume and platelet distribution width are associated with mild cognitive impairment and Alzheimer's disease. J Psychiatr Res 2013 May;47(5):644-649. [doi: 10.1016/j.jpsychires.2013.01.014] [Medline: 23395109]
62. Winchester LM, Powell J, Lovestone S, Nevado-Holgado AJ. Red blood cell indices and anaemia as causative factors for cognitive function deficits and for Alzheimer's disease. Genome Med 2018 Jun 28;10(1):51. [doi: 10.1186/s13073-018-0556-z] [Medline: 29954452]
63. Guo Y, Li P, Ma X, et al. Association of circulating cholesterol level with cognitive function and mild cognitive impairment in the elderly: a community-based population study. Curr Alzheimer Res 2020;17(6):556-565. [doi: 10.2174/1567205017666200810165758] [Medline: 32781960]
64. Magaki S, Yellon SM, Mueller C, Kirsch WM. Immunophenotypes in the circulation of patients with mild cognitive impairment. J Psychiatr Res 2008 Feb;42(3):240-246. [doi: 10.1016/j.jpsychires.2007.01.004] [Medline: 17320906]
65. Yoshida M, Higashi K, Kuni K, et al. Distinguishing mild cognitive impairment from Alzheimer's disease with acrolein metabolites and creatinine in urine. Clinica Chimica Acta 2015 Feb;441:115-121. [doi: 10.1016/j.cca.2014.12.023]
66. Gao H, Schneider S, Hernandez R, et al. Early identification of cognitive impairment in community environments through modeling subtle inconsistencies in questionnaire responses: machine learning model development and validation. JMIR Form Res 2024 Nov 13;8:e54335. [doi: 10.2196/54335] [Medline: 39536306]

67. Graham SA, Lee EE, Jeste DV, et al. Artificial intelligence approaches to predicting and detecting cognitive decline in older adults: a conceptual review. Psychiatry Res 2020 Feb;284:112732. [doi: 10.1016/j.psychres.2019.112732] [Medline: 31978628]
68. Zhou Y, Song Z, Han X, Li H, Tang X. Prediction of Alzheimer's disease progression based on magnetic resonance imaging. ACS Chem Neurosci 2021 Nov 17;12(22):4209-4223. [doi: 10.1021/acschemneuro.1c00472] [Medline: 34723463]

Abbreviations

AD: Alzheimer disease
ADL: activities of daily living
BRF: balanced random forest
CHO: total cholesterol
CLHLS: Chinese Longitudinal Healthy Longevity Survey
HDL: high-density lipoprotein cholesterol
IADL: instrumental activities of daily living
MCI: mild cognitive impairment
MDA: malondialdehyde
ML: machine learning
MMSE: Mini-Mental State Examination
MRI: magnetic resonance imaging
RF: random forest
SHAP: Shapley Additive Explanations
SMOTE: Synthetic Minority Over-Sampling Technique
SVM: support vector machines
XGBoost: Extreme Gradient Boosting

Edited by H Köttl; submitted 19.10.24; peer-reviewed by A Garikipati, Y Wu; revised version received 05.03.25; accepted 01.04.25; published 30.04.25.

Please cite as:
Ren H, Zheng Y, Li C, Jing F, Wang Q, Luo Z, Li D, Liang D, Tang W, Liu L, Cheng W
Using Machine Learning to Predict Cognitive Decline in Older Adults From the Chinese Longitudinal Healthy Longevity Survey: Model Development and Validation Study
JMIR Aging 2025;8:e67437
URL: https://aging.jmir.org/2025/1/e67437
doi:10.2196/67437

The Relationship Between Sleep Disorders and Combination of Diabetes and Sarcopenia in Adults Aged 45 Years or Older: 10-Year Nationwide Prospective Cohort Study

Shugang Li[1], PhD; Yimi Wang[2], MBBS; Linzhi Li[2], MBBS; Hao Wu[1], MSci

[1]School of General Practice and Continuing Education, Capital Medical University, No.10 Xitoutiao, Youanmenwai, Fengtai District, Beijing, China
[2]School of Medicine, Shihezi University, Shihezi, China

Corresponding Author:
Hao Wu, MSci
School of General Practice and Continuing Education, Capital Medical University, No.10 Xitoutiao, Youanmenwai, Fengtai District, Beijing, China

Abstract

Background: With changes in lifestyle, the issue of sleep disorders is becoming increasingly common. Diabetes and sarcopenia have been found to be independently associated with sleep disorders. However, fewer studies have explored the interaction between the combination of diabetes and sarcopenia at different stages and sleep disorders.

Objective: This study aimed to explore the relationship between the combination of diabetes and sarcopenia and the incidence of sleep disorders in adults aged 45 years and older.

Methods: Based on data from the CHARLS (China Health and Retirement Longitudinal Study), we selected participants with comprehensive diagnostic information on diabetes and sarcopenia from 2011 who had normal sleep duration at baseline and checked their follow-up information of sleep duration from 2013, 2015, 2018, and 2020. Diabetes was classified into diabetes (D), prediabetes (PD), and nondiabetes (ND), and sarcopenia was divided into sarcopenia (S), possible sarcopenia (PS), and nonsarcopenia (NS). The participants were divided into DS, DPS, DNS, PDS, PDPS, PDNS, NDS, NDPS, and NDNS groups. Kaplan-Meier survival curves, the log-rank test, Cox proportional hazards regression, and restricted cubic spline models were used for statistical analysis.

Results: A total of 4936 participants were included in this study. The DS group had the highest incidence of sleep disorders: 49.32%, 28.57%, 36.36%, and 80.00% in 2013, 2015, 2018, and 2020 respectively. In the crude model, compared with the NDNS group, the risk of sleep disorders was increased in the DS group (hazard ratio [HR] 1.707, 95% CI 1.196 - 2.437), PDS (HR 1.599, 95% CI 1.235 - 2.071), NDS (HR 1.465, 95% CI 1.282 - 1.674), and DPS group (HR 1.318, 95% CI 1.097 - 1.583). The risk was increased but not statistically significant in the PDPS group (HR 1.160, 95% CI 0.987 - 1.365). After adjusting for covariates, the risk of sleep disorders remained statistically significant in the DS group (HR 1.515, 95% CI 1.059 - 2.167) and was significantly higher in the PDS (HR 1.423, 95% CI 1.096 - 1.847) and NDS (HR 1.279, 95% CI 1.113 - 1.468) groups than that in the NDNS group. The nonlinear associations between appendicular skeletal muscle mass, grip strength, 5-time chair test, fasting plasma glucose, and sleep disorders were observed and described.

Conclusions: The combination of diabetes and sarcopenia significantly increases the risk of sleep disorders in adults aged 45 years and older. and the implementation of progression control of both diabetes and sarcopenia may be helpful to prevent sleep disorders in this population.

(JMIR Aging 2025;8:e66372) doi:10.2196/66372

KEYWORDS

diabetes; sarcopenia; combined effect; sleep disorders; elderly; cohort study

Introduction

With changes in lifestyle, the issue of sleep disorders is becoming increasingly common. Studies report the worldwide prevalence of obstructive sleep apnea, poor sleep quality, other sleep problems, insomnia, and excessive daytime sleepiness to be 46%, 10%, 37%, 29%, and 19% respectively [1]. Even among adolescents, the prevalence of short and disturbed sleep is as high as 34.1% [2]. The impact of sleep disorders on the health of the older adults is more obvious, which can lead to cardiovascular disease [3], dementia [4], depression, and many other health problems. With the acceleration of aging, sleep disorders in older adults has become a global public health problem. A large number of studies have focused on the factors influencing sleep disorders. In addition to psychological [5], environmental [6], lifestyle [7] factors, chronic diseases, such

as diabetes [8] and sarcopenia [9] are found to be closely related to sleep disorders. According to the International Diabetes Federation report, a total of 537 million adults worldwide had diabetes in 2021, which is expected to increase by 46% to 783 million by 2045 [10]. The prevalence of diabetes in people aged ≥45 years in China is 10.6% [11], and there are significant differences between men and women, and individuals in urban and rural areas. With the increase in age, human function significantly declines, and coupled with the impact of metabolic diseases, sarcopenia is becoming more common in older adults. With a reported prevalence of 27% in individuals over 60 years of age [12], sarcopenia has become one of the common comorbidities in older adults.

Diabetes is an important factor affecting sleep disorders [13]. Studies have found a significant association between obstructive sleep apnea risk and type 2 diabetes (odds ratio [OR] 2.44, 95% CI 1.78-3.35; *P*<.001) [14]. Compared with patients without diabetes, both patients with type 1 (hazard ratio [HR] 1.41, 95% CI 1.37-1.46,) and type 2 diabetes (HR 1.40, 95% CI 1.37 - 1.44) have a significantly increased risk of sleep disorders [15]. In addition, studies have shown that sarcopenia is a chronic disease that induces sleep disorders [16], and people with both sarcopenia and diabetes have a significantly increased risk of sleep abnormalities [17]. Even self-reported sleep duration may provide the possibility for early prediction of sarcopenia. These findings suggest a complicated relationship between diabetes, sarcopenia, and sleep disorders.

However, these existing studies have not elucidated the risk of sleep disorders in populations with comorbid diabetes and sarcopenia; most studies have analyzed the relationships between diabetes and sleep disorders, as well as the impact of sarcopenia on sleep disorders. Another important issue is that both diabetes guidelines and the expert consensus on sarcopenia emphasize prediabetes and possible sarcopenia, which are similar to large icebergs beneath the surface—and their contribution to sleep disorders also warrants attention. With aging, the prevalence of diabetes and sarcopenia gradually increases in the older population, which prompts us to consider whether an interaction between diabetes and sarcopenia at different stages significantly increases the risk of sleep disorders. Previous studies have mostly explored the relationship between sarcopenia, diabetes, and sleep disorders using a cross-sectional design. On the one hand, there is a lack of evidence from longitudinal cohort designs; on the other hand, the interaction of the risk between diabetes and sarcopenia at different stages and sleep disorders has not been elucidated. Based on this scientific question, the present study used the China Health and Retirement Longitudinal Study (CHARLS) to analyze the impact of the interaction of diabetes and sarcopenia on the occurrence of sleep disorders over four longitudinal follow-up visits (ie, 2013, 2015, 2018, 2020) from the 2011 baseline population, to provide a reference for the prevention and treatment of sleep disorders in adults aged ≥45 years.

Methods

Participants and Study Design

The participants for this study were derived from the CHARLS, a nationally representative longitudinal survey initiated in 2011. It covered 150 county-level units, 450 village-level units, and 10,257 households across 28 provinces. These participants were followed up every 2 to 3 years. CHARLS covers basic personal information, household structure, health status, physical measurements, health service utilization, health insurance, and basic community characteristics. Four waves of follow-up data collection have been completed so far, and detailed information on CHARLS has been published [18].

Data were collected from CHARLS during 2011 - 2020. All information related to diabetes, sarcopenia, and sleep disorders among participants aged ≥45 years could be downloaded from 2011, 2013, 2015, 2018, and 2020.

We selected participants according to the following steps and excluded based on the following criteria: (1) lack of age information or age<45 years in 2011; 2) missing height and weight measurements at baseline; (3) height and weight data beyond mean (3 SD); (4) missing gender information; (5) missing sarcopenia-related information; (6) missing FPG (fasting plasma glucose) information; (7) FPG <3.9 mmol/L, (8) no sleep status at baseline, (9) nightly sleep duration <6 or >8 hours; (10) questionnaire lacked diabetes self-report information, (11) normal FPG but choosing a self-reported diagnosis of “diabetes or high blood sugar” for the question “Have you been diagnosed with by a doctor?”.

Subsequently, all the participants selected at baseline were matched by ID to the database of 2013, 2015, 2018, and 2020, respectively; those without necessary information about sleep duration were excluded. Only the first record of occurrence of sleep disorder was retained, if one ID was found repeatedly across years.

Measures

The measurements of exposure variables (ie, diabetes, sarcopenia), outcome variables (ie, sleep disorder, change of sleep duration or status), and covariates (ie, age, gender, education level, habitation, current marital status, current smoking, and drinking status) were respectively assessed across the cohort at each survey wave.

Assessment of Diabetes

All participants were classified as having diabetes, prediabetes, and nondiabetes according to the Guidelines for the Prevention and Treatment of type 2 Diabetes Mellitus in China (2020 edition) [19]. Since only FPG in venous blood was measured in the CHARLS, the diagnosis in this study was based only on FPG values. A venous FPG value ≥7.0 mmol/L was classified as diabetes, FPG ≥6.1 mmol/L and <7.0 mmol/L as prediabetes, and FPG <6.0 mmol/L as nondiabetes. Although the baseline survey included the question “Have you been diagnosed with by a doctor?”, the participants answering “diabetes or high blood sugar, (including impaired glucose tolerance and elevated fasting

blood sugar)," were excluded, as it was not possible to distinguish between diabetes and prediabetes.

Assessment of Sarcopenia

The Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment (AWGS 2019) was used to assess the status of the sarcopenia and classify participants into sarcopenia, possible sarcopenia, and nonsarcopenia groups. The average of maximum values of handgrip strength was measured using either or both hands to evaluate the skeletal muscle strength. According to AWGS 2019, the threshold for low handgrip strength is <18 kg for women and <28 kg for men. The 5-time chair stand test was used to evaluate physical performance of the participants; a time ≥12 s of the 5-time chair stand test was defined as low physical performance. Individuals who could not complete the test were excluded. The muscle mass of the participants was evaluated using a previously validated formula [20]:

ASM=0.193×body weight(kg)+0.107×height(cm)−4.157×sex(1=male, 2=female)−0.037×age(years)−2.631

Following the estimation of appendicular skeletal muscle mass index (ASM) using the above formula, the appendicular skeletal muscle mass index (ASMI) was calculated using the ASM divided by the square of height in meters (ASMI = ASM/height^2 m^2). Therefore, the cut-off values for ASMI used in this study were based on the percentile value of the lowest 20% of the study population; low muscle mass was defined as ASMI <7.05 kg/m^2 for men and <5.36 kg/m^2 for women. Sarcopenia was diagnosed if muscle mass, muscle strength, or physical performance were low, and possible sarcopenia was diagnosed if muscle strength or physical performance was low.

Covariates

Based on the literature review, some common factors correlated with sleep disorders, including demographic factors (eg, age and gender), area of residence (ie, city or rural), current marital status (ie, married or unmarried), and education level with a cut-off of junior high school graduation were selected as covariates to divide the participants into two groups. Health–related factors included current smoking and alcohol consumption status.

Statistical Analysis

The categorical variables such as gender, education level, and habitation were expressed as percentages. The χ^2 test was used for comparison between groups and Fisher exact test was used for analyses when the χ^2 test was not appropriate. For quantitative indicators such as age, FPG, and muscle mass index, the mean (SD) was used for normally distributed variables. When the variables were not normally distributed, median with interquartile range was used calculated. The 2-tailed *t* test, ANOVA, or Kruskal-Wallis tests were used for comparison among groups, as appropriate. Kaplan-Meier survival analysis and log-rank test were used to compare the incidence of sleep disorders. Based on the different adjustment covariates, three Cox proportional hazards regression models were used to analyze the combined effect of diabetes and sarcopenia on sleep disorders. The restricted cubic spline (RCS) models were used to analyze the exposure-response and nonlinear relationships between FPG, muscle mass index, grip strength, 5-time chair stand test, and sleep disorders. Microsoft Excel (2021), SPSS software (version 27.0; IBM Corp) and R software (version 4.3.2; R Foundation for Statistical Computing) were used to conduct data collation, statistical testing, and RCS modelling respectively. A $P<.05$ was considered statistically significant.

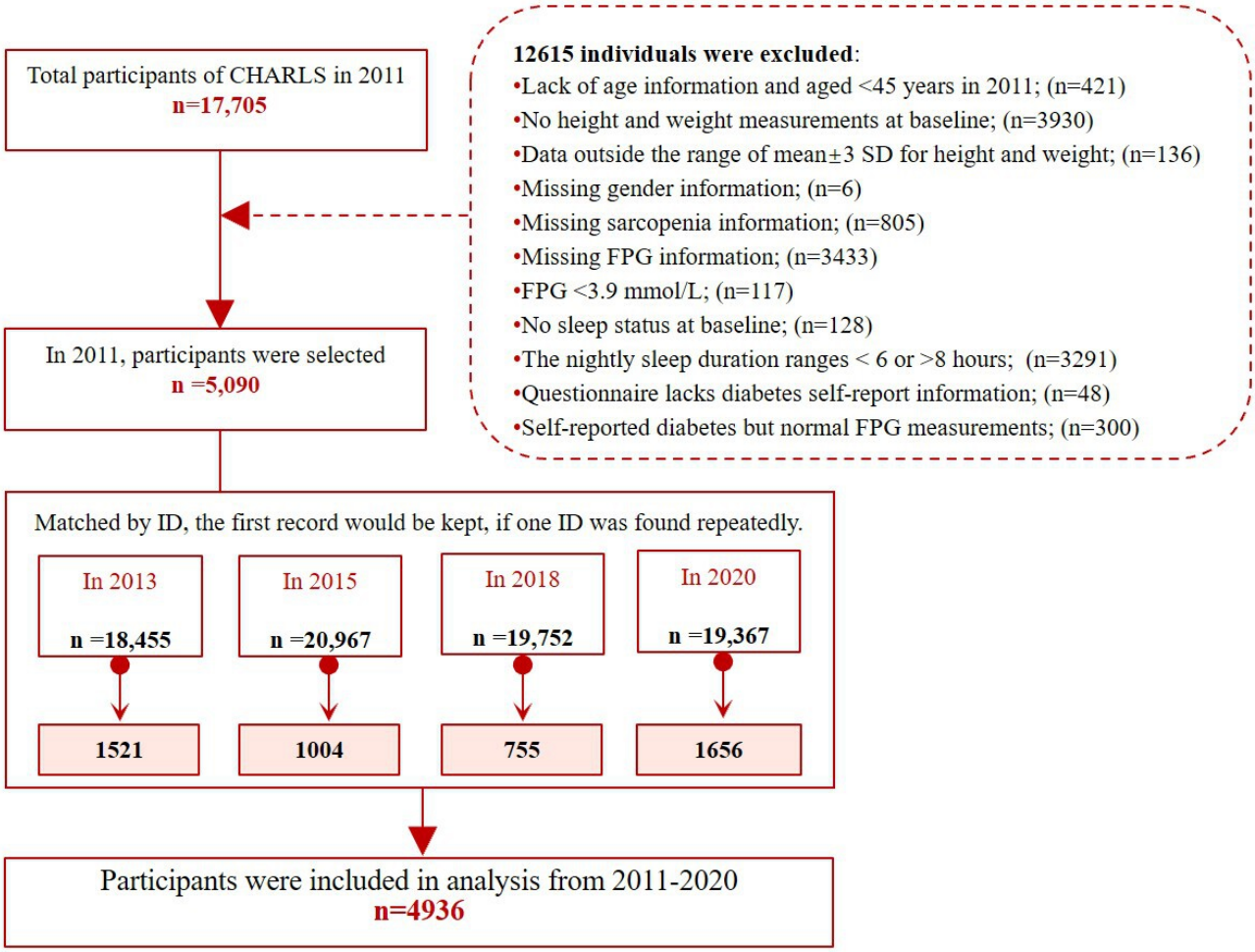
Ethical Considerations

In this study, all data were extracted from CHARLS, which received approval from Peking University Biomedical Ethics Committee (IRB00001052-11015) prior to conducting the original field investigation and all the participants had signed the informed consent form before enrollment. The secondary analysis was approved and anonymized data was used for analysis.

Results

A total of 17,705 participants were included at baseline. Of these, 12,615 were gradually excluded based on the exclusion criteria, and matched with follow-up data from 2013, 2015, 2018, and 2020 according to the participant ID. The first occurrence of sleep disorders was retained in the database. Finally, 4936 participants were included in the analysis (Figure 1).

Figure 1. The flow chart of screening for enrolled individuals. CHARLS: China Health and Retirement Longitudinal Study; FPG: fasting plasma glucose.



A total of 4936 participants were selected and followed up on their sleep during 4 follow-up visits from 2011 to 2020, including 2427 male (49.2%) and 2509 female (50.8%) participants. Further comparison of the combination of diabetes and sarcopenia at different stages—NDNS, NDPS, NDS, PDNS, PDPS, PDS, DNS, DPS, and DS were statistically significant for age, sex, marital status, place of residence, education level, smoking, alcohol consumption, review status, weight, grip strength, skeletal muscle mass, 5-time chair stand test, and blood glucose levels *(*all *Ps*<.05; Table 1).

Table . Sociodemographic characteristics and critical variables of participants.

Variables	NDNS (Nondiabetes and Nonsarcopenia), n=2288	NDPS (Nondiabetes and Possible sarcopenia), n=841	NDS (Nondiabetes and Sarcopenia), n=324	PDNS (Prediabetes and Nonsarcopenia), n=599	PDPS (Prediabetes and Possible sarcopenia), n=230	PDS (Prediabetes and Sarcopenia), n=75	DNS (Diabetes and Nonsarcopenia), n=380	DPS (Diabetes and Possible sarcopenia), n=162	DS (Diabetes and Sarcopenia), n=37	H (*df*)/χ2 (*df*)	*P* value
Age (years), median (IQR)	55 (49-61)	58 (52-64)	68 (60-74)	57 (50-63)	60 (55-65)	68 (60-75)	57 (51-63)	59 (55-65)	70 (62-75)	528.833[a] (8)	<.001
Gender, n (%)										37.412[b] (8)	<.001
Male	1168 (51.0)	349 (41.5)	146 (45.1)	323 (53.9)	111 (48.3)	39 (52.0)	203 (53.4)	68 (42.0)	20 (54.1)		
Female	1120 (49.0)	492 (58.5)	178 (54.9)	276 (46.1)	119 (51.7)	36 (48.0)	177 (46.6)	94 (58.0)	17 (45.9)		
Marriage, n (%)											<.001[c]
Yes	2120 (92.7)	772 (91.8)	261 (80.6)	556 (92.8)	198 (86.1)	58 (77.3)	355 (93.4)	143 (88.3)	28 (75.7)		
No	168 (7.3)	69 (8.2)	63 (19.4)	43 (7.2)	32 (13.9)	17 (22.7)	25 (6.6)	19 (11.7)	9 (24.3)		
Habitation, n (%)											<.001[c]
City	2098 (91.7)	796 (94.6)	314 (96.9)	543 (90.7)	205 (89.1)	73 (97.3)	341 (90.0)	149 (92.0)	35 (94.6)		
Rural	190 (8.3)	45 (5.4)	10 (3.1)	56 (9.3)	25 (10.9)	2 (2.7)	38 (10.0)	13 (8.0)	2 (5.4)		
Education, n (%)										169.139[b] (8)	<.001
Above of junior high school graduation	946 (41.3)	225 (26.8)	45 (13.9)	235 (39.2)	68 (29.6)	12 (16.0)	159 (42.0)	40 (24.7)	6 (16.2)		
Others	1342 (58.7)	616 (73.2)	278 (86.1)	364 (60.8)	162 (70.4)	63 (84.0)	220 (58.0)	122 (75.3)	31 (83.8)		
Smoking	935 (40.9)	298 (35.5)	117 (36.1)	277 (46.2)	90 (39.1)	38 (50.7)	160 (42.1)	62 (38.3)	17 (45.9)	24.306[b] (8)	.002
Drinking	864 (37.8)	234 (27.8)	91 (28.1)	248 (41.4)	72 (31.3)	24 (32.0)	165 (43.4)	41 (25.3)	9 (24.3)		<.001
Height(cm), median (IQR)	159.5(153.4-165.5)	157.8(152.4-163.9)	154.9 (146.8-161.0)	159.5 (153.3-165.8)	158.5 (153.5-164.8)	154.0 (146.0-154.0)	159.5 (152.9-165.3)	156.0(151.7-162.3)	157.0(151.0-163.2)	138.643[a] (8)	<.001
Weight (kg), median (IQR)	58.7 (52.1-66.2)	60.2 (55.1-67.0)	45.6 (42.2-49.3)	61.0 (54.6-69.4)	63.3 (57.3-69.2)	46.6 (42.3-50.3)	61.6 (53.3-69.9)	61.4 (55.8-68.7)	46.4 (42.5-51.2)	878.573[a] (8)	<.001
Handgrip strength (kg), median (IQR)	34.0 (28.0-41.0)	26.5 (20.0-34.4)	22.5 (17.0-27.7)	34.0 (28.5-41.6)	27.5 (19.3-35.2)	22.6(17.5-29.5)	34.0 (28.1-41.3)	25.0 (18.5-32.5)	21.6 (17.0-29.3)	818.091[a] (8)	<.001
ASM[d] (kg/m^2), median (IQR)	17.8 (14.3-20.7)	16.7 (14.4-20.3)	12.9 (10.6-17.1)	18.1 (14.9-21.5)	18.2 (14.9-21.1)	13.9 (10.6-17.7)	18.3 (14.6-21.3)	16.7 (14.4-20.4)	14.7 (11.6-17.3)	343.836[a] (8)	<.001

Variables	NDNS (Nondiabetes and Nonsarcopenia), n=2288	NDPS (Nondiabetes and Possible sarcopenia), n=841	NDS (Nondiabetes and Sarcopenia), n=324	PDNS (Prediabetes and Nonsarcopenia), n=599	PDPS (Prediabetes and Possible sarcopenia), n=230	PDS (Prediabetes and Sarcopenia), n=75	DNS (Diabetes and Nonsarcopenia), n=380	DPS (Diabetes and Possible sarcopenia), n=162	DS (Diabetes and Sarcopenia), n=37	H (*df*)/χ2 (*df*)	*P* value
5-time chair stand test (s), median (IQR)	8.6 (7.2-10.2)	13.5 (12.3-15.4)	13.3 (11.2-15.5)	8.9 (7.4-10.3)	13.5 (12.2-16.3)	13.3 (11.3-16.3)	9.0 (7.5-10.3)	13.0 (12.3-15.4)	13.6 (11.2-17.3)	2336.735[a] (8)	<.001
FPG[e] (mmol/L), median (IQR)	5.5 (5.2-5.8)	5.5 (5.1-5.8)	5.4 (5.1-5.7)	6.4 (6.3-6.6)	6.4 (6.2-6.6)	6.4 (6.2-6.6)	8.0 (7.3-9.5)	8.0 (7.4-9.4)	7.8 (7.3-8.9)	3208.235[a] (8)	<.001

[a] Kruskal-Wallis H test.

[b]χ^2: chi-square statistic.

[c]Fisher's exact test.

[d]ASM: appendicular skeletal muscle mass.

[e]FPG: fasting plasma glucose.

Based on survival analysis, it was observed that with the change in follow-up time, there was a significant difference in the incidence of sleep disorders among all groups (χ^{28}=73.486, *P*<.001; Figure 2). In 2020, the DS group had the highest incidence of sleep disorders (80. and the incidence in the DPS, PDS and NDS groups reached 52.00%, 50.00%, and 48.78%, respectively; whereas, the incidence of sleep disorders was the lowest in the NDNS group (36.02%). Furthermore, the incidence of sleep disorders in the DS group was 49.32%, 28.57%, 36.36%, and 80.00% in 2013, 2015, 2018, and 2020, respectively, showing a significantly increasing trend. Significant differences were also observed between men and women, and the incidence of short sleep duration (<6 hours) and long sleep duration (>8 hours), as shown in Figure 3.

Figure 2. The incidence of sleep disorders of the co-exposure of diabetes and sarcopenia at different stages. In the diabetes group, (+) represents diabetes, (±) represents prediabetes, and (-) depicts nondiabetes. In the sarcopenia group, (+) represents sarcopenia, (±) indicates possible sarcopenia, and (-) depicts nonsarcopenia.

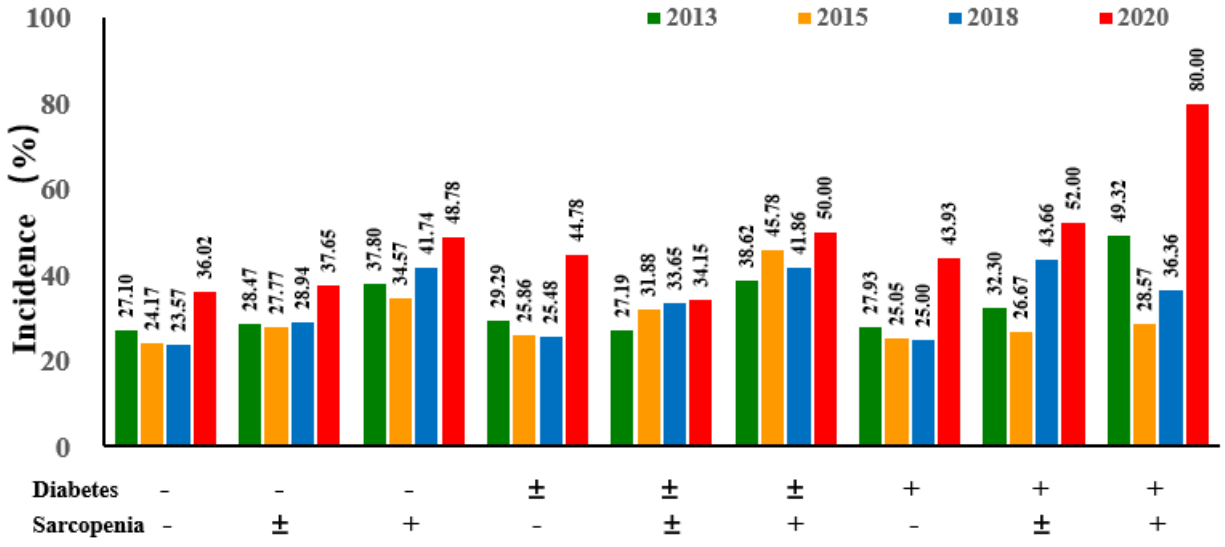
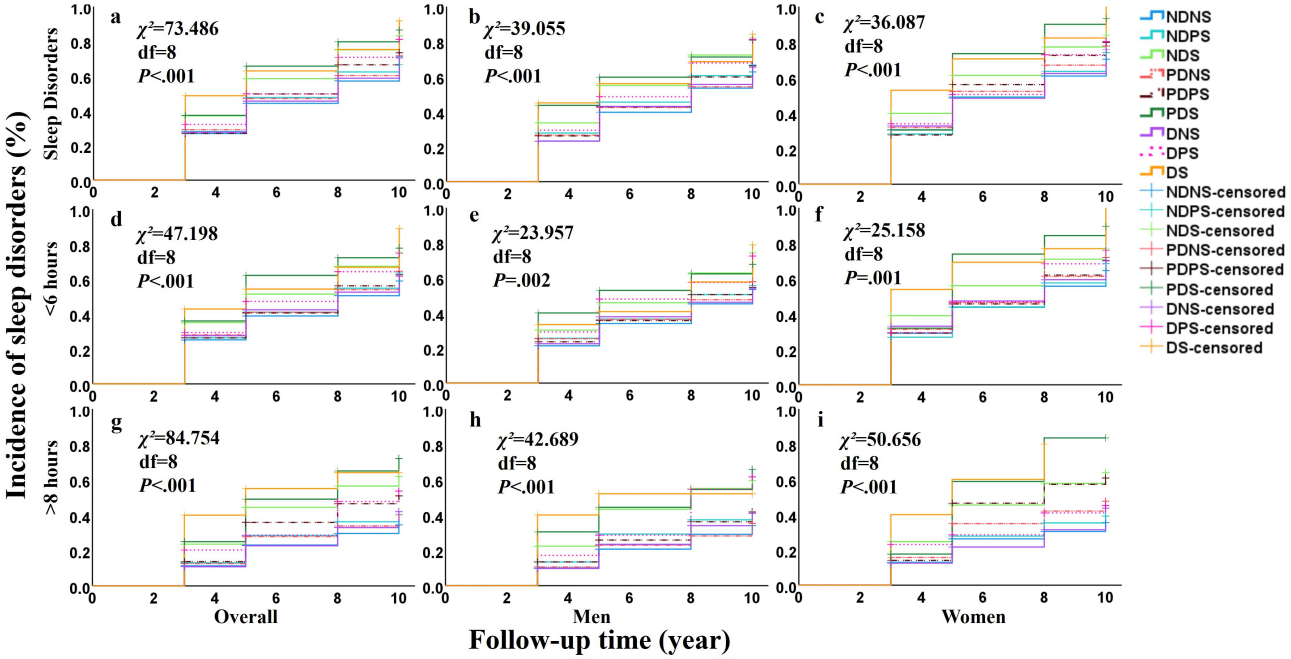


Figure 3. Incidence of sleep disorders related to co-exposure of diabetes and sarcopenia among groups. Follow-up time depicts the number of years of follow-up in 2013, 2015, 2018, 2020, respectively. χ^2 indicates the statistic from the log-rank test. NDNS: Nondiabetes and Nonsarcopenia, NDPS: Nondiabetes and Possible Sarcopenia, NDS: Nondiabetes and Sarcopenia, PDNS: Prediabetes and Nonsarcopenia, PDPS: Prediabetes and Possible Sarcopenia, PDS: Prediabetes and Sarcopenia, DNS: Diabetes and Nonsarcopenia, DPS: Diabetes and Possible Sarcopenia, DS: Diabetes and Sarcopenia; SD: sleep disorders (including sleep duration <6h and >8h); <6 hours: sleep duration at the final follow-up was <6 hours; >8 hours: sleep duration at the final follow-up was >8 hours.



According to Cox regression analysis, there was a significant association between the combined status of diabetes mellitus, sarcopenia, and sleep disorders. In Model 1, no covariates were adjusted, compared with the NDNS group, the risk of sleep disorders was increased in the DS group (HR 1.707, 95% CI 1.196 - 2.437, *P*=.003), PDS (HR 1.599, 95% CI 1.235 - 2.071, *P*<.001), NDS (HR 1.465, 95% CI 1.282 - 1.674, *P*<.001), and DPS (HR 1.318, 95% CI 1.097 - 1.583, *P*=.003). The risk was increased but not statistically significant in the PDPS group. After adjusting for the variables (Model 3), the risk of sleep disorders in the DS group (HR 1.515, 95% CI 1.059 - 2.167; *P*=.02) was still statistically significant, and the risk in the PDS (HR 1.423, 95% CI 1.096 - 1.847; *P*=.008) and NDS (HR 1.279,95% CI 1.113 - 1.468; *P*<.001) groups were significantly higher than that in the NDNS group (Table 2).

Table . Association between sleep disorders and co-exposure to diabetes and sarcopenia at different status. Symbols indicate existing condition: diabetes (+), nondiabetes (–), prediabetes (±) in the diabetes group, and sarcopenia (+), possible sarcopenia (–), and no sarcopenia (±) in the sarcopenia group.

Group		Model 1[a], HR[b] (95% CI)	Model 2[c], HR (95% CI)	Model 3[d], HR (95% CI)
Diabetes	Sarcopenia			
Sleep disorders[e]				
-	-	Reference	Reference	Reference
-	±	1.105 (1.003-1.218)	1.058 (0.959-1.167)	1.033 (0.936-1.139)
-	+	1.465 (1.282-1.674)	1.344 (1.171-1.541)	1.279 (1.113-1.468)
±	-	1.107 (0.992-1.235)	1.100 (0.986-1.277)	1.098 (0.984-1.225)
±	±	1.160 (0.987-1.365)	1.114 (0.946-1.311)	1.110 (0.943-1.307)
±	+	1.599 (1.235-2.071)	1.501 (1.157-1.946)	1.423 (1.096-1.847)
+	-	1.067 (0.935-1.219)	1.061 (0.929-1.212)	1.070 (0.936-1.223)
+	±	1.318 (1.097-1.583)	1.251 (1.040-1.504)	1.228 (1.021-1.477)
+	+	1.707 (1.196-2.437)	1.594 (1.115-2.279)	1.515 (1.059-2.167)
Sleep duration <6 hours				
-	-	Reference	Reference	Reference
-	±	1.100 (0.981-1.234)	1.046 (0.932-1.174)	1.023 (0.911-1.149)
-	+	1.455 (1.233-1.716)	1.343 (1.134-1.591)	1.286 (1.084-1.525)
±	-	1.120 (0.985-1.273)	1.120 (0.985-1.273)	1.118 (0.983-1.271)
±	±	1.086 (0.884-1.334)	1.059 (0.861-1.302)	1.061 (0.863-1.305)
±	+	1.596 (1.139-2.235)	1.503 (1.071-2.110)	1.417 (1.008-1.992)
+	-	1.073 (0.915-1.258)	1.070 (0.913-1.255)	1.075 (0.916-1.261)
+	±	1.366 (1.099-1.697)	1.287 (1.035-1.601)	1.260 (1.012-1.568)
+	+	1.713 (1.123-2.612)	1.623 (1.062-2.481)	1.546 (1.010-2.365)
Sleep duration >8 hours				
-	-	Reference	Reference	Reference
-	±	1.208 (1.007-1.450)	1.137 (0.946-1.367)	1.073 (0.892-1.290)
-	+	2.132 (1.696-2.681)	1.791 (1.415-2.268)	1.601 (1.262-2.030)
±	-	1.197 (0.972-1.474)	1.173 (0.952-1.445)	1.160 (0.942-1.429)
±	±	1.594 (1.222-2.080)	1.465 (1.121-1.914)	1.429 (1.093-1.869)
±	+	2.545 (1.699-3.814)	2.256 (1.502-3.389)	2.033 (1.351-3.058)
+	-	1.198 (0.941-1.525)	1.162 (0.913-1.480)	1.195 (0.937-1.524)
+	±	1.665 (1.184-2.343)	1.516 (1.076-2.136)	1.502 (1.065-2.118)
+	+	2.739 (1.414-5.306)	2.340 (1.205-4.543)	1.970 (1.011-3.840)
Decreased sleep duration[f]				
-	-	Reference	Reference	Reference
-	±	1.122 (1.014-1.242)	1.073 (0.969-1.189)	1.052 (0.949-1.166)
-	+	1.368 (1.175-1.593)	1.274 (1.090-1.490)	1.224 (1.046-1.433)
±	-	1.076 (0.958-1.208)	1.069 (0.952-1.201)	1.067 (0.950-1.199)
±	±	1.121 (0.933-1.349)	1.091 (0.907-1.313)	1.101 (0.914-1.325)
±	+	1.735 (1.261-2.386)	1.617 (1.174-2.229)	1.525 (1.105-2.104)
+	-	1.063 (0.923-1.224)	1.058 (0.918-1.219)	1.069 (0.928-1.232)
+	±	1.251 (1.018-1.537)	1.188 (0.966-1.461)	1.167 (0.948-1.435)
+	+	1.546 (1.015-2.357)	1.452 (0.951-2.216)	1.391 (0.910-2.124)

Group		Model 1[a], HR[b] (95% CI)	Model 2[c], HR (95% CI)	Model 3[d], HR (95% CI)
Diabetes	Sarcopenia			
Increased sleep duration[g]				
-	-	Reference	Reference	Reference
-	±	1.253 (1.078-1.457)	1.212 (1.042-1.409)	1.167 (1.002-1.358)
-	+	1.674 (1.363-2.056)	1.468 (1.188-1.814)	1.385 (1.119-1.713)
±	-	1.115 (0.937-1.325)	1.099 (0.924-1.308)	1.081 (0.908-1.286)
±	±	1.377 (1.094-1.733)	1.295 (1.028-1.631)	1.259 (0.998-1.588)
±	+	1.952 (1.367-2.787)	1.815 (1.270-2.595)	1.697 (1.185-2.429)
+	-	1.104 (0.901-1.353)	1.078 (0.879-1.322)	1.088 (0.886-1.335)
+	±	1.270 (0.941-1.714)	1.196 (0.885-1.616)	1.201 (0.888-1.624)
+	+	2.067 (1.167-3.662)	1.852 (1.043-3.288)	1.641 (0.920-2.924)

[a]Model 1 represents no covariates were adjusted.

[b]HR: hazard ratio.

[c]Model 2 represents adjusting for gender and age.

[d]Model 3 represents adjustment for gender, age, education, current marital status and habitation, smoking, and drinking.

[e]Sleep disorders including sleep duration <6 hours and >8 hours.

[f]Decreased indicates difference of sleep duration between baseline and follow-up <0.

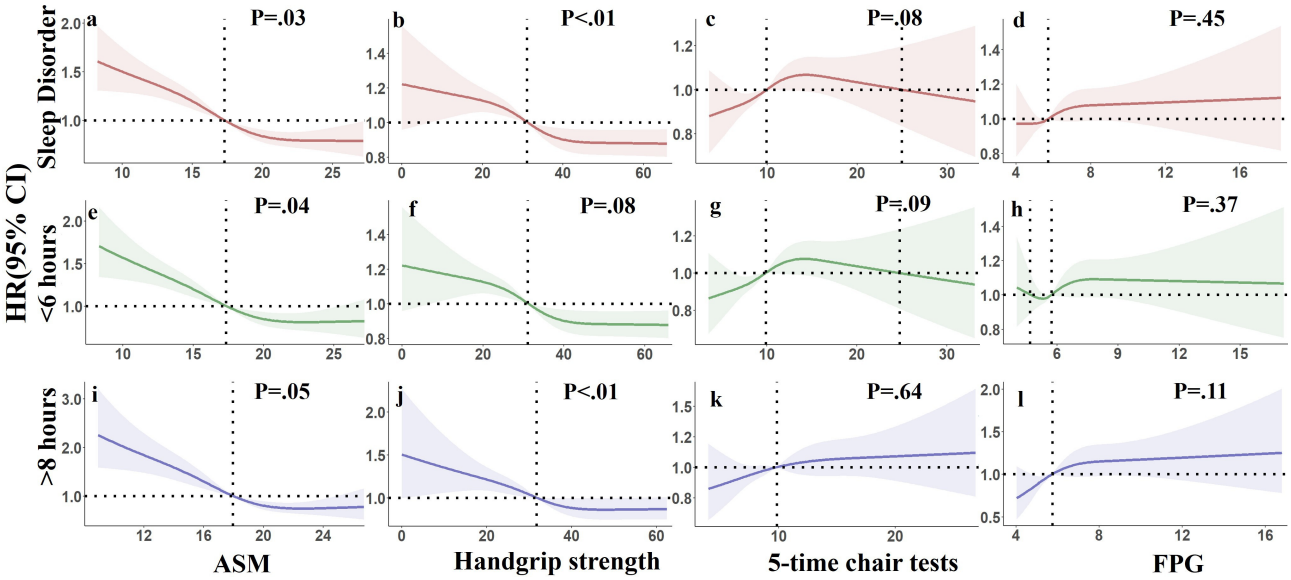
[g]Increased indicates the difference of sleep duration between baseline and follow-up >0.

After a stratified analysis of sleep duration (ie, <6 hours and >8 hours), the risk of short sleep duration (< 6 hours) was significantly increased in the DS group without adjusting for variables (Model 1: HR 1.713, 95% CI 1.123 - 2.612) and remained statistically significant after adjusting for variables (Model 3: HR 1.546, 95% CI 1.010 - 2.365). The risk of long sleep duration (> 8 hours) was significantly increased in the DS group without adjusting for variables (Model 1: HR 2.739, 95% CI 1.414 - 5.306), which remained statistically significant after adjusting for variables (Model 3: HR 1.970, 95% CI 1.011 - 3.840; Table 2).

To analyze the changes in sleep duration, we stratified participants based on the difference in sleep duration between the final follow-up and baseline. If the difference was <0, it was defined as “Decrease,” while a difference >0, was defined as “Increase in sleep duration.” Cox regression analysis revealed that the risk of “Decrease” in sleep duration in the NDS group without adjusting for the variable (Model 1) is significantly higher (HR 1.735, 95% CI 1.261 - 2.386), and this remained statistically significant after adjusting for covariates (Model 3: (HR 1.525, 95% CI 1.105 - 2.104). After adjusting for covariates, the risk of “Increase” in the PDS (Model 3: HR 1.697, 95% CI 1.185 - 2.429) and NDS (Model 3: HR 1.385, 95% CI 1.119 - 1.713) groups also increased significantly (Table 2).

To further assess the relationship between diabetes, sarcopenia, and sleep disorders, the RCS model was used to describe important factors related to sarcopenia—ASM, handgrip strength, 5-time chair stand test performance—and the exposure response relationship between FPG levels, and sleep disorders. Figure 4 illustrates these association of sleep disorders with ASM (χ^2=45.04, P<.001, P for nonlinearity=.0284), handgrip strength (χ^2=16.62, P<.008, P for nonlinearity<.0017), and 5-time chair stand tests (χ^2=8.78, P=.03, P for nonlinearity=.0806). A nonlinear correlation was observed between sleep disorders and both ASM and handgrip strength. The relationship between FPG and long sleep duration (>8 hours) was also significant (χ^2=9.44, P=.02, P for nonlinearity=.10731).

Figure 4. Exposure-response relationship among diabetes, sarcopenia, and sleep disorders.. The restricted cubic spline (RCS) was used to model expose-response curves after adjusting for factors including gender, age, education, current marital status, and habitation, smoking, and drinking. ASM: appendicular skeletal muscle mass; FPG: fasting plasma glucose; HR: hazard ratio; SD: sleep disorders; <6 h: sleep duration was <6 hours; >8 h: sleep duration was >8 hours.



Discussion

With increasing age, the issue of sleep disorders in older adults is becoming more severe [21]. Common chronic conditions including diabetes mellitus and sarcopenia which are highly prevalent could affect sleep quality. In particular, prediabetes and possible sarcopenia are now being diagnosed at an early age. If this status is not addressed during reversible stages, the risk of sleep disorders in adults aged ≥45 years is even more severe. Although studies have reported a relationship between diabetes [22] or sarcopenia [23] and sleep disorders, there is a lack of studies on the relationship between co-exposure at varying stages of diabetes and sarcopenia. This study found that participants with diabetes and sarcopenia had the highest risk of sleep disorders; groups with prediabetes and sarcopenia, nondiabetes with sarcopenia, and diabetes with possible sarcopenia had a significantly higher risk of sleep disorders. Additionally, there was an exposure-response relationship between ASM, handgrip strength, 5-time chair test, FPG levels, and sleep disorders. These findings can provide a reference for the prevention of sleep disorders in adults aged ≥45 years and the development of early screening and personalized intervention.

Diabetes, as a chronic disease with high incidence in older adults, is an important factor contributing to sarcopenia [24] and sleep disorders [25]. Our findings showed that participant groups with both diabetes and prediabetes have an increased risk of sleep disorders. This may be explained by the following mechanisms. On one hand, diabetes can affect the function of the hypothalamic-pituitary-adrenal axis, resulting in abnormal secretion of hormones such as cortisol [26]. The rhythm of these hormones can affect the sleep-wake cycle, leading to the sleep disorders [27]. On the other hand, hyperglycemia may lead to the secretion of melatonin, which plays a key role in regulating sleep rhythms [28]. Its altered secretion can lead to difficulty falling asleep and poor sleep quality [29]. In addition, patients with sarcopenia are prone to muscle fatigue and damage due to the loss of mass and strength of several muscles [30]. During sleep, changes in body posture may strain already fragile muscles, causing pain and discomfort, resulting in sleep disruption. Muscle strength in the chest and abdomen decreases in patients with sarcopenia, resulting in decreased respiratory function [31]. During sleep, weakness in respiratory muscles may cause poor breathing, snoring and even apnea, thus interfering with sleep [32]. It is also important to note that diabetes can cause sarcopenia. Long-term hyperglycemia may lead to peripheral neuropathy [33], causing a decrease in muscle blood supply, affecting the nutrition and innervation of nerves to muscles [34], which in turn leads to muscle waste atrophy. Patients with diabetes and prediabetes have widespread insulin resistance, which blocks insulin signaling pathways such as PI3K/Ak [35] and Ras-MAPK [36], thus cannot effectively promote the uptake of amino acids by muscle cells, thereby inhibiting the synthesis of muscle proteins; this results in a decrease in muscle mass. These mechanisms reveal the relationship between FPG, ASM, handgrip strength, 5-time chair stand test, and sleep disorders, as identified by the RCS model.

This study found that participants with comorbid diabetes and sarcopenia had the highest risk of sleep disorders, and the groups with prediabetes and sarcopenia, nondiabetes and sarcopenia, and diabetes with possible sarcopenia also had a significantly higher risk of sleep disorders. The potential mechanisms of sleep disorders caused by the comorbidity of diabetes and sarcopenia may be related to the following: (1) patients with diabetes already have abnormal blood sugar regulation; sarcopenia will further affect the uptake and utilization of glucose by muscles, resulting in more frequent and difficult-to-control blood sugar fluctuations [37]. This may result in the superposition of metabolic disorders, triggering energy metabolism imbalance [38], making the body unable to obtain enough energy at night to maintain normal physiological

functions and sleep status. (2) Neuropathy caused by diabetes, including damage to peripheral and autonomic nerves can affect the sensory and motor functions of muscles [39]. Sarcopenia aggravates muscle weakness and atrophy, further impairing the function of the neuromuscular junction [40]. This neuromuscular dysfunction can lead to nocturnal muscle twitches, spasms, or pain that interfere with sleep. Autonomic dysfunction is common in both diabetes and sarcopenia; it affects the regulation of the cardiovascular and respiratory systems, resulting in abnormal fluctuations in blood pressure and heart rate at night, poor breathing, and increases the risk of sleep disruption [41]. Both diabetes and sarcopenia are associated with abnormal hormone secretion [42], which may lead to reduced secretion of testosterone, estrogen, and insulin-like growth factor 1 (IGF-1) among others. Significant gender differences were observed in our study, which may further support this phenomenon. Furthermore, diabetes and sarcopenia are both accompanied by chronic inflammation and increased levels of oxidative stress [43]. Tumor necrosis factor α (TNF-α), interleukin 6 (IL-6) promote muscle catabolism and may affect the central nervous system function, resulting in disorders in sleep regulation centers. Reactive oxygen species produced during oxidative stress can damage both neural and muscular systems, aggravate physical discomfort, and then affect sleep [44].

Our findings have several implications. First, special attention should be paid to adults aged ≥45 years with sleep disorders. To improve sleep duration and quality, it is necessary to effectively manage diabetes, perform physical exercises to improve muscle mass and function, and continuously monitor sleep quality. Second, for people with prediabetes or possible sarcopenia, early intervention should be carried out to prevent the occurrence of diabetes or sarcopenia, which can effectively reduce the risk of sleep disorders. Third, the community should build appropriate sports facilities and provide an environment that promotes physical exercise to improve the health of residents. In addition, regular screening for diabetes and sarcopenia should be implemented to detect patients with diabetes and sarcopenia at an early stage, and comprehensive interventions should be actively used to prevent sleep disorders.

The advantages of this study include the following aspects. It is a large-scale, nationally representative cohort study of community residents, using survival analysis and Cox regression models to explore the co-exposure of diabetes and sarcopenia—as well as the impact of prediabetes with latent risk and possible sarcopenia—on sleep duration and sleep status. However, there are also a few limitations within our study: First, due to limitations in the CHARLS data, we could not fully explore the relationship between diabetes status, progression trends in sarcopenia from 2011 to 2020, and specific types of sleep disorders. Our analysis was limited to examining the impact of diabetes and sarcopenia status on sleep disorders in 2011; the relationship between the factors affecting progression and sleep disorders is complex. Concurrently, due to limited data, we included a smaller number of participants aged ≥75 years in our study, and the differences in population distribution across age groups may limit the extrapolation of our findings. Second, prediabetes is an intermediate stage that can revert toward normoglycemia and diabetes. Therefore, our study only analyzed the effect of prediabetes exposure at baseline on sleep disorders. Since the CHARLS did not provide data on FPG levels for the four follow-up waves , it was impossible to explore the role of changes in prediabetes status on sleep disorders. However, our findings still suggest that patients with prediabetes were at an increased risk of sleep disorders. Therefore, family-based and both individual health education [45], and community-level behavioral interventions play an important role in the prevention of sleep disorders [46]. Fourth, there is a mutual relationship among diabetes, sarcopenia, and sleep disorders. Although both diabetes and sarcopenia can lead to sleep disorders, the interaction between diabetes and sarcopenia on sleep disorders could not be analyzed due to lack of sufficient data. Fifth, this study found a potential relationship between the co-exposure to various diabetes and sarcopenia statuses, and sleep disorders. Diabetes and sarcopenia may have a synergistic effect on the occurrence of sleep disorders; however, causality cannot be based solely on our findings. Future research should focus on the progression and changes of diabetes, sarcopenia, and sleep disorders, including sleep efficiency [47] to explore the relationship among them.

In conclusion, this study found that diabetes combined with sarcopenia in adults aged 45 years and older may increase the risk of sleep disorders. To reduce this risk, it is essential to implement effective disease management for both diabetes and sarcopenia, and to further evaluate joint intervention measures for individuals with prediabetes or possible sarcopenia.

Acknowledgments

This study was funded by the Beijing High level Public Health Technical Talents Training Plan (2022-1-005, Key DisciplineMember-02- ‐ 44). We are grateful to the CHARLS (China Health and Retirement Longitudinal Study), which provided data for this research. We also thank all the participants, staff members, volunteers devoted to the CHARLS project.

Authors' Contributions

Conceptualization: HW, SL
Data curation: SL, YW
Formal analysis: SL, YW
Funding acquisition: HW, SL
Investigation: SL, YW
Methodology: SL, YW

Supervision: HW, SL
Validation: LL
Writing – original draft: SL, YW
Writing – review & editing: HW, LL, SL, YW

Conflicts of Interest

All authors declare no conflicts of interest.

References

1. Canever JB, Zurman G, Vogel F, et al. Worldwide prevalence of sleep problems in community-dwelling older adults: a systematic review and meta-analysis. Sleep Med 2024 Jul;119:118-134. [doi: 10.1016/j.sleep.2024.03.040] [Medline: 38669835]
2. Yue L, Cui N, Liu Z, Jia C, Liu X. Patterns of sleep problems and internalizing and externalizing problems among Chinese adolescents: a latent class analysis. Sleep Med 2022 Jul;95:47-54. [doi: 10.1016/j.sleep.2022.04.008] [Medline: 35561474]
3. Liu SH, Lin FJ, Kao YH, et al. Chronic partial sleep deprivation increased the incidence of atrial fibrillation by promoting pulmonary vein and atrial arrhythmogenesis in a rodent model. Int J Mol Sci 2024 Jul 11;25(14):7619. [doi: 10.3390/ijms25147619] [Medline: 39062858]
4. Robbins R, Quan SF, Weaver MD, Bormes G, Barger LK, Czeisler CA. Examining sleep deficiency and disturbance and their risk for incident dementia and all-cause mortality in older adults across 5 years in the United States. Aging (Albany NY) 2021 Feb 11;13(3):3254-3268. [doi: 10.18632/aging.202591] [Medline: 33570509]
5. Haddadi A, Matinnia N, Yazdi-Ravandi S. The relationship between corona disease anxiety and sleep disturbances and suicidal ideation in medical staff: The mediating role of resiliency and cognitive flexibility: a cross-sectional study. Health Sci Rep 2024 Aug;7(8):e2282. [doi: 10.1002/hsr2.2282] [Medline: 39144407]
6. Wang J, Gueye-Ndiaye S, Castro-Diehl C, et al. Associations between indoor fine particulate matter (PM2.5) and sleep-disordered breathing in an urban sample of school-aged children. Sleep Health 2024 Aug 1:S2352-7218(24)00133-5. [doi: 10.1016/j.sleh.2024.06.004] [Medline: 39095254]
7. Zheng YB, Huang YT, Gong YM, et al. Association of lifestyle with sleep health in general population in China: a cross-sectional study. Transl Psychiatry 2024 Aug 4;14(1):320. [doi: 10.1038/s41398-024-03002-x] [Medline: 39098892]
8. Algethami A, Alfahmi FK, Alhusayni MA, et al. Evaluation of sleep quality among people living with type 2 diabetes mellitus in Taif, Saudi Arabia: a cross-sectional study. Cureus 2024 Jul;16(7):e64124. [doi: 10.7759/cureus.64124] [Medline: 39119415]
9. Liu K, Luo J, Chen Y, et al. Association between sarcopenia and sleep disorders: a cross-sectional population based study. Front Nutr 2024;11:1415743. [doi: 10.3389/fnut.2024.1415743]
10. Diabetes around the world in 2021. International Diabetes Federation. URL: https://diabetesatlas.org/ [accessed 2025-08-05]
11. Lingling G, Lin Y, Jiao Y. Prevalence and influencing factors of diabetes for the 45 years and older population in China. Modern Preventive Medicine 2019;46:2890-2894 [FREE Full text]
12. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
13. Woo J, Lehrer HM, Tabibi D, Cebulske L, Tanaka H, Steinhardt M. The association of multidimensional sleep health with hba1c and depressive symptoms in African American Adults with type 2 diabetes. Psychosom Med 2024 May 1;86(4):307-314. [doi: 10.1097/PSY.0000000000001298] [Medline: 38724038]
14. Taimah M, Ahmad A, Al-Houqani M, et al. Association between obstructive sleep apnea risk and type 2 diabetes among Emirati adults: results from the UAE healthy future study. Front Endocrinol (Lausanne) 2024;15:1395886. [doi: 10.3389/fendo.2024.1395886] [Medline: 39081790]
15. Ni MH, Yang YS, Huang JY, Lo SC, Huang CN, Kornelius E. The association of depression and sleep disorders in patients with type 1 diabetes in Taiwan. Medicine (Baltimore) 2024 Jul 19;103(29):e38969. [doi: 10.1097/MD.0000000000038969] [Medline: 39029051]
16. Lv X, Peng W, Jia B, Lin P, Yang Z. Longitudinal association of sleep duration with possible sarcopenia: evidence from CHARLS. BMJ Open 2024 Mar 23;14(3):e079237. [doi: 10.1136/bmjopen-2023-079237] [Medline: 38521528]
17. Ida S, Kaneko R, Nagata H, et al. Association between sarcopenia and sleep disorder in older patients with diabetes. Geriatr Gerontol Int 2019 May;19(5):399-403. [doi: 10.1111/ggi.13627] [Medline: 30773802]
18. Zhao Y, Strauss J, Yang G, Giles J. 2011-2012 national baseline user's guide [chinese]. China Health and Retirement Longitudinal Study. 2013. URL: https://charls.charlsdata.com/Public/ashelf/public/uploads/document/2011-charls-wave1/application/Chinese_users__guide_20130407_.pdf [accessed 2024-07-20]
19. Chinese Diabetes Society, Zhu D. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition) [Chinese]. Chin J Diabetes Mellitus 2021;13(4):315-409 [FREE Full text] [doi: 10.3969/j.issn.1006-6187.2022.01.002]
20. Wen X, Wang M, Jiang CM, Zhang YM. Anthropometric equation for estimation of appendicular skeletal muscle mass in Chinese adults. Asia Pac J Clin Nutr 2011;20(4):551-556. [Medline: 22094840]

21. Delbari A, Ahmadi F, Zar A, Zandvakili A, Sadeghipour HR, Sims J. Living in urban or rural environments affect the sleep quality of the elderly in Bushehr (Southern Iran): emphasizing the active and inactive of the elderly. BMC Public Health 2024 May 18;24(1):1340. [doi: 10.1186/s12889-024-18747-9] [Medline: 38760757]
22. Tenda ED, Henrina J, Cha JH, et al. Obstructive sleep apnea: overlooked comorbidity in patients with diabetes. World J Diabetes 2024 Jul 15;15(7):1448-1460. [doi: 10.4239/wjd.v15.i7.1448] [Medline: 39099813]
23. Smith L, Shin JI, Veronese N, et al. Sleep duration and sarcopenia in adults aged ≥ 65 years from low and middle-income countries. Aging Clin Exp Res 2022 Jul;34(7):1573-1581. [doi: 10.1007/s40520-022-02074-3] [Medline: 35103953]
24. Chowdhury SR, Chakrabarti A, Datta PK. Sarcopenia in patients with diabetes mellitus, an overlooked perioperative condition. Br J Anaesth 2023 Sep;131(3):e65-e67. [doi: 10.1016/j.bja.2023.06.032]
25. Kuroda H, Yeung SLA, Fujii R, Iwagami M, Goto A. Investigating the non-linear association between sleep duration and type 2 diabetes: conventional and Mendelian randomization analyses from the UK Biobank. J Diabetes Investig 2025 Apr 3:40181521. [doi: 10.1111/jdi.70039] [Medline: 40181521]
26. Marissal-Arvy N, Moisan MP. Diabetes and associated cognitive disorders: role of the hypothalamic-pituitary adrenal axis. Metabol Open 2022 Sep;15:100202. [doi: 10.1016/j.metop.2022.100202] [Medline: 35958117]
27. Ba-Ali S, Brøndsted AE, Andersen HU, Sander B, Jennum PJ, Lund-Andersen H. Assessment of diurnal melatonin, cortisol, activity, and sleep-wake cycle in patients with and without diabetic retinopathy. Sleep Med 2019 Feb;54:35-42. [doi: 10.1016/j.sleep.2018.10.018] [Medline: 30529775]
28. Onaolapo AY, Adebisi EO, Adeleye AE, Olofinnade AT, Onaolapo OJ. Dietary melatonin protects against behavioural, metabolic, oxidative, and organ morphological changes in mice that are fed high-fat, high-sugar diet. Endocr Metab Immune Disord Drug Targets 2020;20(4):570-583. [doi: 10.2174/1871530319666191009161228] [Medline: 32138638]
29. van Maanen A, Meijer AM, Smits MG, van der Heijden KB, Oort FJ. Effects of melatonin and bright light treatment in childhood chronic sleep onset insomnia with late melatonin onset: a randomized controlled study. Sleep 2017 Feb 1;40(2):28364493. [doi: 10.1093/sleep/zsw038] [Medline: 28364493]
30. Zhang Y, Chen X, Hou L, et al. Prevalence and risk factors governing the loss of muscle function in elderly sarcopenia patients: a longitudinal study in China with 4 years of follow-up. J Nutr Health Aging 2020;24(5):518-524. [doi: 10.1007/s12603-020-1361-0] [Medline: 32346691]
31. Nagano A, Wakabayashi H, Maeda K, et al. Respiratory sarcopenia and sarcopenic respiratory disability: concepts, diagnosis, and treatment. J Nutr Health Aging 2021;25(4):507-515. [doi: 10.1007/s12603-021-1587-5] [Medline: 33786569]
32. Peng X, Zhou R, Liu C, Chen X, Zhu T, Chen G. Abnormal sleep duration is associated with sarcopenia in older Chinese people: A large retrospective cross-sectional study. Open Med (Wars) 2024;19(1):20240938. [doi: 10.1515/med-2024-0938] [Medline: 38584821]
33. Albariqi MM, Versteeg S, Brakkee EM, et al. Human IAPP is a contributor to painful diabetic peripheral neuropathy. J Clin Invest 2023 Apr 17;133(8):e156993. [doi: 10.1172/JCI156993] [Medline: 36917177]
34. van Sloten TT, Savelberg HHCM, Duimel-Peeters IGP, et al. Peripheral neuropathy, decreased muscle strength and obesity are strongly associated with walking in persons with type 2 diabetes without manifest mobility limitations. Diabetes Res Clin Pract 2011 Jan;91(1):32-39. [doi: 10.1016/j.diabres.2010.09.030] [Medline: 20965601]
35. Feng Y, Ren Y, Zhang X, et al. Metabolites of traditional Chinese medicine targeting PI3K/AKT signaling pathway for hypoglycemic effect in type 2 diabetes. Front Pharmacol 2024;15:1373711. [doi: 10.3389/fphar.2024.1373711]
36. Kasowanjete P, Abrahamse H, Houreld NN. Photobiomodulation at 660 nm stimulates in vitro diabetic wound healing via the Ras/MAPK Pathway. Cells 2023 Apr 4;12(7):1080. [doi: 10.3390/cells12071080] [Medline: 37048153]
37. Norlin S, Axelsson J, Ericsson M, Edlund H. O304 ameliorates hyperglycemia in mice by dually promoting muscle glucose effectiveness and preserving β-cell function. Commun Biol 2023 Aug 25;6(1):877. [doi: 10.1038/s42003-023-05255-6] [Medline: 37626210]
38. Al-Sayyar A, Hammad MM, Williams MR, Al-Onaizi M, Abubaker J, Alzaid F. Neurotransmitters in type 2 diabetes and the control of systemic and central energy balance. Metabolites 2023 Mar 4;13(3):384. [doi: 10.3390/metabo13030384] [Medline: 36984824]
39. Zilliox LA. Diabetes and peripheral nerve disease. Clin Geriatr Med 2021 May;37(2):253-267. [doi: 10.1016/j.cger.2020.12.001] [Medline: 33858608]
40. Miao Y, Xie L, Song J, et al. Unraveling the causes of sarcopenia: roles of neuromuscular junction impairment and mitochondrial dysfunction. Physiol Rep 2024 Jan;12(1):e15917. [doi: 10.14814/phy2.15917] [Medline: 38225199]
41. Mizukami K. Autonomic dysfunction in dementia with Lewy bodies: focusing on cardiovascular and respiratory dysfunction. PCN Rep 2023 Sep;2(3):e129. [doi: 10.1002/pcn5.129] [Medline: 38867816]
42. He Y, Fu Q, Sun M, et al. Phosphoproteome reveals molecular mechanisms of aberrant rhythm in neurotransmitter-mediated islet hormone secretion in diabetic mice. Clin Transl Med 2022 Jun;12(6):e890. [doi: 10.1002/ctm2.890] [Medline: 35758323]
43. Muvhulawa N, Mazibuko-Mbeje SE, Ndwandwe D, et al. Sarcopenia in a type 2 diabetic state: reviewing literature on the pathological consequences of oxidative stress and inflammation beyond the neutralizing effect of intracellular antioxidants. Life Sci 2023 Nov 1;332:122125. [doi: 10.1016/j.lfs.2023.122125] [Medline: 37769808]
44. Singh G, Singh K, Sinha RA, Singh A, Kumar A. Japanese encephalitis virus infection causes reactive oxygen species - mediated skeletal muscle damage. Eur J of Neuroscience 2024 Sep;60(5):4843-4860. [doi: 10.1111/ejn.16469]

45. Amerzadeh M, Shafiei Kisomi Z, Senmar M, Khatooni M, Hosseinkhani Z, Bahrami M. Self-care behaviors, medication adherence status, and associated factors among elderly individuals with type 2 diabetes. Sci Rep 2024 Aug 18;14(1):19118. [doi: 10.1038/s41598-024-70000-w] [Medline: 39155329]
46. Van Ancum JM, Meskers CGM, Reijnierse EM, et al. Lack of knowledge contrasts the willingness to counteract sarcopenia among community-dwelling adults. J Aging Health 2020;32(7-8):787-794. [doi: 10.1177/0898264319852840] [Medline: 31156027]
47. Sakal C, Li T, Li J, Yang C, Li X. Association between sleep efficiency variability and cognition among older adults: cross-sectional accelerometer study. JMIR Aging 2024 Apr 4;7:e54353. [doi: 10.2196/54353] [Medline: 38596863]

Abbreviations

ASM: appendicular skeletal muscle mass
ASMI: appendicular skeletal muscle mass index
AWGS: Asian Working Group for Sarcopenia: 2019 Consensus Update on Sarcopenia Diagnosis and Treatment
CHARLS: China Health and Retirement Longitudinal Study
FPG: fasting plasma glucose
HR: hazard ratio
RCS: restricted cubic spline

Edited by H Köttl; submitted 11.09.24; peer-reviewed by WQ Chen, Z Yang; revised version received 13.04.25; accepted 17.04.25; published 02.06.25.

Please cite as:
Li S, Wang Y, Li L, Wu H
The Relationship Between Sleep Disorders and Combination of Diabetes and Sarcopenia in Adults Aged 45 Years or Older: 10-Year Nationwide Prospective Cohort Study
JMIR Aging 2025;8:e66372
URL: https://aging.jmir.org/2025/1/e66372
doi:10.2196/66372

Sarcopenia and Risk of Cognitive Impairment: Cohort Study and Mendelian Randomization Analysis

Tingting Sha[1,2,3], PhD; Yuqing Zhang[4,5], DSc; Jie Wei[1,2,3,6,7], PhD; Changjun Li[3,6,8], MD, PhD; Chao Zeng[1,2,3,6,7], MD, PhD; Guanghua Lei[1,2,3,6], MD, PhD; Yilun Wang[1,2,3], MD, PhD

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

[2]Hunan Key Laboratory of Joint Degeneration and Injury, Changsha, China

[3]Key Laboratory of Aging-related Bone and Joint Diseases Prevention and Treatment of Ministry of Education, Changsha, China

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

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

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

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

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

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

Abstract

Background: Over half the people over 60 years of age experience cognitive impairment, with limited treatment options, making it crucial to identify risk factors. Studies have examined the association between sarcopenia and cognitive impairment; however, the evidence is inconclusive and cannot be used to make causal inferences.

Objective: This study aims to appraise the causal association of sarcopenia with cognitive impairment by triangulating the data from a cohort study and Mendelian randomization (MR) analysis.

Methods: Using UK Biobank data, we first examined the associations of sarcopenia and its indices (appendicular lean mass [ALM], handgrip strength, and gait speed) with cognitive function (fluid intelligence and prospective memory) by using mixed-effects regression models. Then, we explored the causal associations of genetically predicted sarcopenic indices with cognitive function through a 2-sample MR, and examined potential mediation by omega-3 fatty acids, vitamin D levels, physical inactivity, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes.

Results: A total of 34,457 participants, with a mean age of 56.4 (SD 7.6) years, 51.1% (n=17,620) of which were female, completed baseline cognitive tests between 2006 and 2010 and attended at least 1 follow-up visit in 2012, 2014, or 2019, and were included in the observational analysis. The cohort study revealed that sarcopenia was significantly associated with cognitive impairment, which was evidenced by reduced fluid intelligence scores (β=−0.91, 95% CI −1.68 to −0.15; *P*=.02). Each of the sarcopenic indices also exhibited significant associations with either fluid intelligence or prospective memory (all *P*<.05). MR analyses yielded compelling evidence of positive associations between the genetically predicted increases in ALM (β=0.09, 95% CI 0.07-0.12; *P*<.001), handgrip strength (β=0.18, 95% CI 0.08-0.29; *P*<.001) and gait speed (β=0.78, 95% CI 0.53-0.29; *P*<.001) and improved cognitive function. The effects of ALM and handgrip strength on cognitive function were partially mediated by genetically predicted physical activity, with indirect effects of 0.01 (95% CI 0.00-0.02) for ALM and 0.02 (95% CI 0.00-0.05) for handgrip strength.

Conclusions: Our study suggests that sarcopenia is a potential causal risk factor for cognitive impairment, with physical activity acting as a modifiable mediator in this relationship.

(JMIR Aging 2025;8:e66031) doi:10.2196/66031

KEYWORDS

sarcopenia; musculoskeletal disease; cognitive impairment; cognition; mediation; Mendelian randomization; genetic variation; cohort study

Introduction

An impairment in cognitive function is a neurodegenerative process that affects 10% - 20% of people aged 65 years and older [1], and that can lead to adverse health consequences, including diminished quality of life [2] and increased risk of hospitalization [3], as well as mortality [4]. However, by the time of diagnosis, the pathological changes related to cognitive impairment have often become irreversible. The available treatments are limited and target symptoms, with very low efficacy [5]. Thus, identifying the potential factors for predicting the risk of subsequent cognitive impairment has become a public health priority [6,7]. Such measurable indices can help recognize high-risk populations to explore potential disease-modifying therapies.

Sarcopenia, characterized by accelerated loss of muscle mass and deterioration in function, is a prevalent skeletal muscle disorder in the elderly [8]. As predicted by statistics, about 2 billion individuals globally will be affected by sarcopenia by 2050 [9]. Through the progression of physical inactivity, patients with sarcopenia tend to experience an increased risk of disability, which has been associated with cognitive impairment due to its impact on neurogenesis and cerebral blood vessel formation [10,11]. Previous cohort studies have reported a possible relationship between sarcopenia and cognitive impairment [12-17]. However, this relationship is still not fully understood and may even be controversial [18]. Some studies have even rejected any significant association between sarcopenia itself [19], or its related indices [20,21] and cognitive impairment. Discrepancies in definitions of sarcopenia and sample sizes among studies may underlie these conflicting outcomes. Hence, a more intricate exploration is warranted to establish the precise link between sarcopenia and cognitive impairment.

Furthermore, current available human evidence on such association is mostly based on observational research, and therefore cannot be relied on to derive causal inferences due to inherent limitations (eg, reverse causality and unmeasured confounders) [22]. Examining the causal association between sarcopenia and cognitive impairment and identifying potentially modifiable intervention targets along the causal pathway is of important public health significance for preventing cognitive impairment [18]. Mendelian randomization (MR) is a method that involves using genetic variants as proxies for the targeted exposure. Since, genetic variants are conditional on parental genotypes and randomly allocated at conception, the results from MR studies are more resistant to reverse causality and confounding than those derived from conventional observational studies [22].

Using a large population-based cohort, this study was initiated by evaluating the observational data that links the consensus definition of sarcopenia, which was established by the European Working Group on Sarcopenia in Older People in 2019 (EWGSOP2), as well as its 3 defining indices (ie, appendicular lean mass [ALM], handgrip strength, and gait speed), to the risk of cognitive impairment. To best assess causality, we performed multiple MR analyses to explore the potential causal association between sarcopenia and cognitive function. Subsequently, we further used MR mediation analysis to examine the degree to which 11 putative mediators (ie, omega-3 fatty acids, vitamin D levels, physical inactivity, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes) may impact the effects of sarcopenia, to identify potential intervention targets in the causal pathway. This study aims to elucidate the causal relationship between sarcopenia and cognitive impairment while identifying actionable intervention targets within the causal pathway through the evaluation of these mediators.

Methods

Study Design

We examined the associations of sarcopenia and its indices with cognitive function in a cohort of approximately 35,000 participants from the United Kingdom Biobank. We then conducted multiple MR analyses to explore the potential causal relations between sarcopenic indices (ie, ALM, handgrip strength, and gait speed) and cognitive function. Finally, we used mediation MR analysis to quantify to what extent physical inactivity, depression, anxiety, falls, frailty, and vitamin D use mediated the effects of sarcopenic indices on cognitive function. This research has been conducted using the UK Biobank Resource under Application 77,646.

Cohort Study

Study Participants

The UK Biobank is a comprehensive prospective cohort that enrolled ≥500,000 participants aged 40 - 69 from 22 centers over the United Kingdom between 2006 and 2010. Neuropsychological tests were performed to assess cognitive function assessment in a subset of participants at baseline as well as 3 subsequent visits via a touch screen. Approximately 170,000 individuals participated in the baseline visit (2006 - 2010), and ~20,000, ~60,000, and ~8000 individuals participated in the first (2012 - 2013), second (2014+), and third (2019+) follow-up visits, respectively. For each follow-up visit dataset, consistent diagnostic criteria were applied for both sarcopenia and cognitive function. In our analysis, we included 35,000 participants who underwent cognitive function tests at baseline and had at least one subsequent follow-up visit (see Figure S1 in Multimedia Appendix 1).

Exposure and Outcome Measurements

Sarcopenia was defined according to the EWGSOP2, which uses the detection of low grip strength and low muscle mass to confirm sarcopenia, with poor physical performance indicating severe sarcopenia [23]. ALM was evaluated by bioelectrical impedance analysis with a Tanita BC418MA body composition analyzer and expressed in kilograms. We used the cut points as recommended in the EWGSOP2 definition of <7 kg/m^2 in men and 5.5 kg/m^2 in women. Grip strength was measured with a hydraulic handheld dynamometer (Jamar J00105). Specifically, 3 measurements were recorded for the maximum strength of both hands, and the highest values were used for analysis, expressed in kilograms. The cut-points for low grip strength

recommended by EWGSOP2 were <27 kg in men and <16 kg in women. Gait speed was acquired directly from the participants through the question "How would you characterize your usual walking speed?" (options: "slow" [<3 miles per hour], "steady/average" [3 - 4 miles per hour], "fast" [>4 miles per hour], or "prefer not to answer" [regarded as missing data]). Participants who reported that they were unable to walk or walked at a slow pace had low physical performance.

Cognitive function was assessed through 2 neuropsychological tests, namely fluid intelligence and prospective memory, by using a touch screen during visits to the UK Biobank assessment center. The fluid intelligence test was designed as a 13-item problem-solving task aiming at logic and reasoning abilities. Prospective memory, which is a form of episodic memory, was used as an indicator to measure the participants' capacity to remember future tasks. The participants were instructed to reproduce a figure on a touch screen following a single instruction to be recalled later in the session. The responses were judged as either "correct on the first attempt" or otherwise, indicating a lapse in prospective memory (eg, "instruction not remembered, either skipped or incorrect" or "correctly recalled on the second attempt").

Covariate Measurements

The selection of potential confounders was based on previous literature [12,13], which was assessed through a touch-screen questionnaire at baseline. These variables included: age (continuous), sex (male or female), education level (college or university; A levels (Advanced Level), AS level (Advanced Subsidiary Level) or equivalent; O level, GCSEs (General Certificate of Secondary Education) or equivalent; CSEs or equivalent; NVQ (National Vocational Qualification), HND (Higher National Diploma), HNC (Higher National Certificate) or equivalent; other professional qualifications), self-reported race (White, Asian, Black, Mixed or other), assessment center, Townsend Deprivation Index at recruitment (continuous), BMI (continuous), presence of long-standing illness (no or yes), overall health rating (excellent, good, fair or poor), smoking status (never, former, or current), frequency of alcohol intake (daily or almost daily; 3 or 4 times a week; once or twice a week; 1 to 3 times a month or special occasions only), sleep duration (<7 h/day, 7 - 9 h/day, or >9 h/day), and TV viewing duration (continuous). Specifically, the presence of long-standing illness was assessed using the following question, "Do you have any long-standing illness, disability, or infirmity". Overall health rating was collected by asking "In general how would you rate your overall health?" All categorical variables strictly adhered to UK Biobank classification standards, with no additional category merging or processing.

Statistical Analyses

To address the repeated measurements of the 2 cognitive tests, we used mixed-effects linear and logistic regression models to evaluate the associations of sarcopenia and its indices with fluid intelligence scores and with the risk of prospective memory impairment, respectively. All the mixed-effects models featured a random individual-specific intercept with a fixed slope. In regression models, we adjusted for sociodemographic, socioeconomic factors, and the number of cognitive assessments, and then further added health-related factors, including BMI, long-standing illness, and health rating. Finally, lifestyle factors, including smoking, alcohol intake frequency, sleep duration, and duration of TV viewing were added. We used a missing indicator approach in the primary analysis and conducted a sensitivity analysis by excluding missing data to evaluate its impact.

MR Analyses

Data Source of Sarcopenic Indices and General Cognitive Function

Given the lack of genome-wide association studies (GWASs) specifically focused on sarcopenia, we used the publicly available summary data on sarcopenic indices and cognitive function from the UK Biobank [24,25] and the COGENT Consortium [26]. The GWAS summary data for ALM (n=450,243), handgrip strength (n=461,089), and gait speed (n=461,089) were derived from the largest public GWASs of individuals with European ancestry in the UK Biobank [24,25]. The summary genetic statistics for the associations of general cognitive function were extracted from a comprehensive GWAS conducted through the UK Biobank and COGENT Consortium [26]. The summary phenotype source and outcome descriptions are provided in Table S1 in Multimedia Appendix 1.

Genetic Instruments for Sarcopenic Indices

The single nucleotide polymorphism (SNPs) meeting the following criteria were selected as instrumental variables: (1) significantly associated with ALM, handgrip strength, or usual gait speed ($P<5\times10^{-8}$), (2) independent (linkage disequilibrium $r^2<.001$ within 10,000 kb), (3) with appropriate effect allele frequencies (≥1%), and (4) not palindromic (adenine/thymine or cytosine/guanine). To correct for multiple comparisons (3 exposures), the Bonferroni method was used. Associations with $P<.016$ (0.05/3) were deemed statistically significant.

Data Analyses

The 2-Sample MR Analyses

The associations between genetically predicted sarcopenic indices and general cognitive function were examined by multiplicative random-effects inverse variance weighted (IVW) analysis, which can provide the most accurate and unbiased estimates [27]. Furthermore, we performed MR-Egger method, MR-PRESSO method, and RadialMR method to pinpoint potential violations of MR assumptions and assess the robustness of primary results [27,28]. The RadialMR method identifies outliers influencing MR analysis, and the results are reanalyzed after their removal [29]. A consistent estimate across multiple sensitivity analyses indicates strengthened causal evidence. We also assessed the bias and type 1 error rate for sample overlap using an internet-based calculator [30]. In addition, we reanalyzed the data using the MRlap method, which is robust to biases caused by sample overlap, winner's curse, and weak instruments [31].

Mediation MR Analyses

We performed univariable MR to estimate the effects of sarcopenic indices on 11 genetically predicted putative

mediators, ie, omega-3 fatty acids, vitamin D levels, moderate-to-vigorous intensity physical activity (MVPA) during leisure time, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, and type 2 diabetes. Mediators showing causal evidence were selected for multivariable Mendelian randomization (MVMR) analysis to estimate the indirect effect of sarcopenic indices on cognitive function mediated by each. The mediation proportion was calculated as the indirect effect divided by the total effect on cognitive function, with standard errors estimated by the delta method [32]. If an inconsistent mediation was observed, where the direct effect opposes the indirect effect, no mediation proportion would be estimated [33].

Complementary Analyses

We performed bidirectional MR analysis to partially explore the potential reverse causality. The SNPs from GWAS that were significantly associated with general cognitive function ($P<5\times10^{-8}$) were selected as instrumental variables (selection criteria remained consistent with the ones mentioned earlier). In addition, we used the MR-Steiger method to examine the directionality of the relationship [34].

All analyses were conducted using the R (R Foundation for Statistical Computing) packages *TwoSampleMR* (version 0.5.6), *MVMR* (version 0.3), *MRPRESSO* (version 1.0), and *MRlap* (version 0.0.3.0) in R (version 4.3.0). A P value <.05 was considered significant. IVW estimates were deemed causal if consistent with at least one sensitivity analysis and showed no pleiotropy (Egger intercept P>.05). Results were reported as odds ratios (ORs), β coefficients, or proportions with 95% confidence intervals (CIs).

Ethical Considerations

This study had been granted with the UK Biobank research approval by the North West Centre for Research Ethics Committee (11/NW/0382) and informed consent was obtained from all participants. For GWAS datasets, ethical review and approval can be accessed in the original studies. The data used were anonymized to ensure privacy and confidentiality. No compensation was provided to participants.

Results

Figure 1 provides an overview of the study design. The baseline characteristics of the cohort study are summarized in Table 1. Among 34,457 participants, 17,620 (51.1%) were women, with a mean age of 56.4 (SD 7.6) years.

Figure 1. Study design flowchart. ALM: appendicular lean mass; MR: Mendelian randomization.

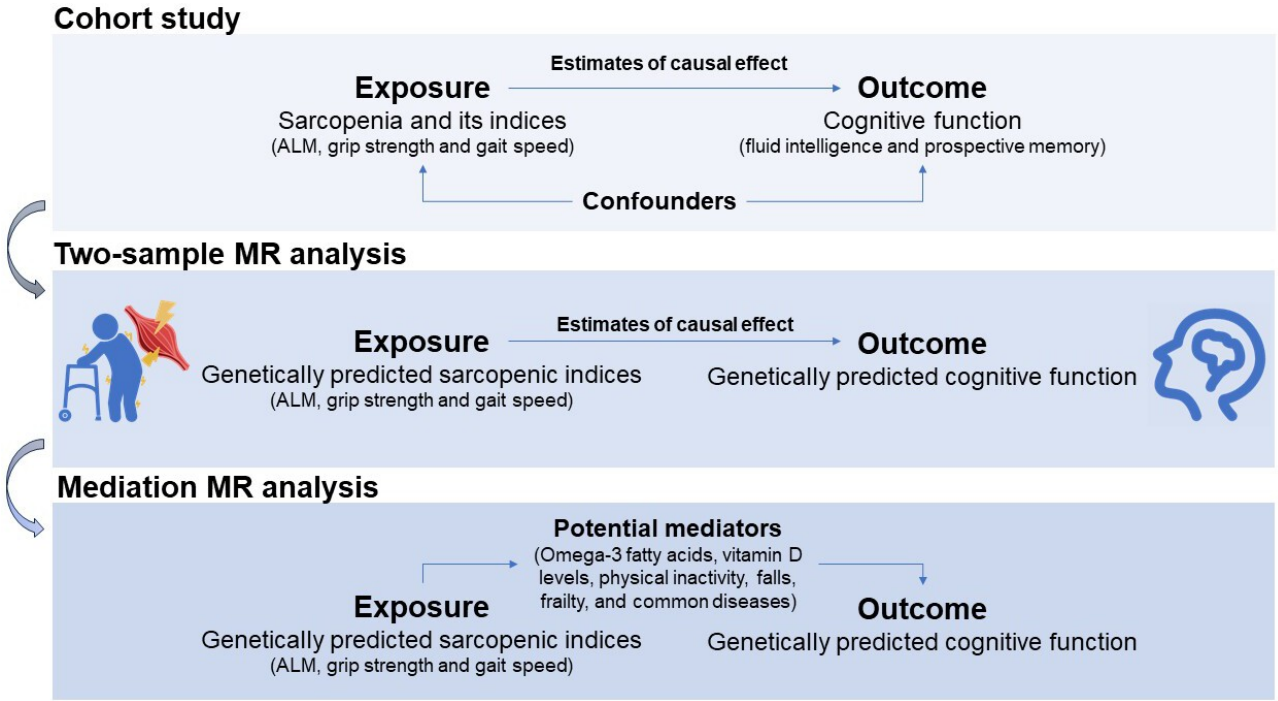


Table . Baseline characteristics of the cohort study.

Characteristics	Results
Sociodemographics	
Total sample, n	34,457
Age (years), mean (SD)	56.4 (7.6)
Gender (female), n (%)	17,620 (51.1)
Deprivation index, mean (SD)	−1.92 (2.7)
Race, n (%)	
White	33,286 (96.6)
Asian	466 (1.4)
Black	240 (0.7)
Mixed or other	162 (0.5)
Unknown	303 (0.9)
Education level, n (%)	
College or university degree	15,190 (44.1)
A[a] levels, AS[b] levels, or equivalent	4355 (12.6)
O levels, GCSEs,[c] or equivalent	6712 (19.5)
CSEs[d] or equivalent	1345 (3.9)
NVQ[e], HND[f], HNC,[g] or equivalent	1985 (5.8)
Other professional qualifications	1747 (5.0)
Unknown	3123 (9.1)
Health-related factors	
BMI, n (%)	
Underweight	158 (0.5)
Normal weight	12,792 (37.1)
Overweight	14,684 (46.2)
Obesity	6750 (19.6)
Unknown	73 (0.2)
Long-standing illness, n (%)	
No	24,258 (70.4)
Yes	9571 (27.8)
Unknown	628 (1.8)
Overall health rating, n (%)	
Excellent	7082 (20.6)
Good	20,969 (60.9)
Fair	5541 (16.1)
Poor	796 (2.3)
Unknown	69 (0.2)
Lifestyle behaviors	
Smoking status, n (%)	
Never	20,501 (59.5)
Former	11718 (34.1)
Current	2163 (6.3)

Characteristics	Results
Unknown	75 (0.2)
Alcohol intake frequency, n (%)	
Daily or almost daily	7768 (22.5)
Three or four times a week	9408 (27.3)
Once or twice a week	8694 (25.2)
One to three times a month	3755 (10.9)
Special occasions only	3069 (8.9)
Never	1749 (5.1)
Unknown	14 (0.0)
Sleep duration, n (%)	
Short sleep (<7 h/day)	7504 (21.8)
Normal (7 - 9 h/day)	26,465 (76.8)
Long sleep (>9 h/day)	386 (1.1)
Unknown	102 (0.3)
TV viewing (h/days) mean (SD)	1.8 (3.0)

[a]A: Advanced.
[b]AS: Advanced Subsidiary.
[c]GCSE: General Certificate of Secondary Education
[d]CSE: Certificate of Secondary Education.
[e]NVQ: National Vocational Qualification.
[f]HND: Higher National Diploma.
[g]HNC: Higher National Certificate.

Cohort Study

As shown in Table 2, participants with sarcopenia had lower fluid intelligence scores than those without sarcopenia, and the multivariable-adjusted difference in the mean level of fluid intelligence scores was 0.91 (95% CI −1.68 to −0.15; *P*=.02). However, no statistically significant association with prospective memory loss was detected (OR 1.36, 95% CI 0.31-5.92; *P*=.68). Each 5-kg increase in ALM was found to be associated with an increased fluid intelligence score (β=0.27, 95% CI 0.21-0.32; *P*<.001) and a decreased risk of prospective memory loss (OR 0.87, 95% CI 0.78-0.97; *P*<.001). Likewise, each 5-kg increase in handgrip strength was positively associated with fluid intelligence (β=0.02, 95% CI 0.01-0.04; *P*<.001) and negatively associated with the prospective memory loss (OR 0.89, 95% CI 0.86-0.93; *P*<.001). Furthermore, slow gait speed was associated with a lower fluid intelligence score (β=−0.10, 95% CI −0.23 to 0.03; *P*=.15) and an increased risk of prospective memory loss (OR 1.65, 95% CI 1.23-2.19; *P*<.001). The results of sensitivity analyses, after excluding cases with missing data, were consistent with the overall analyses (see Table S2 in Multimedia Appendix 1).

Table . The associations of baseline sarcopenia and its indices with follow-up cognitive function.

Sarcopenia indices	Cognitive function	
	Fluid intelligence, β (95% CI)	Prospective memory loss, OR[a] (95% CI)
Sarcopenia		
Model 1[b]	−0.87 (−1.62 to −0.12)[c]	1.67 (0.47-1.78)
Model 2[d]	−0.86 (−1.61 to −0.11)[c]	1.59 (0.39-6.50)
Model 3[e]	−0.91 (−1.68 to −0.15)[c]	1.36 (0.31-5.92)
Appendicular lean mass		
Model 1[b]	0.17 (0.13-0.21)[c]	0.87 (0.80-0.94)[c]
Model 2[d]	0.27 (0.22-0.32)[c]	0.82 (0.74-0.91)[c]
Model 3[e]	0.27 (0.21-0.32)[c]	0.87 (0.78-0.97)[c]
Handgrip strength		
Model 1[b]	0.03 (0.02-0.05)[c]	0.93 (0.89-0.96)[c]
Model 2[d]	0.03 (0.01-0.04)[c]	0.91 (0.88-0.94)[c]
Model 3[e]	0.02 (0.01-0.04)[c]	0.89 (0.86-0.93)[c]
Slow gait speed		
Model 1[b]	−0.20 (−0.32 to −0.08)[c]	1.64 (1.27-2.12)[c]
Model 2[d]	−0.14 (−0.27 to −0.01)[c]	1.71 (1.30-2.25)[c]
Model 3[e]	−0.10 (−0.23 to 0.03)	1.65 (1.23-2.19)[c]

[a]OR: odds ratio.

[b]Model 1 was adjusted for age, sex, race, aged race, education, deprivation index, and the number of cognitive assessments.

[c]*P*<.05.

[d]Model 2 was additionally adjusted for health-related factors including BMI and long-standing illness.

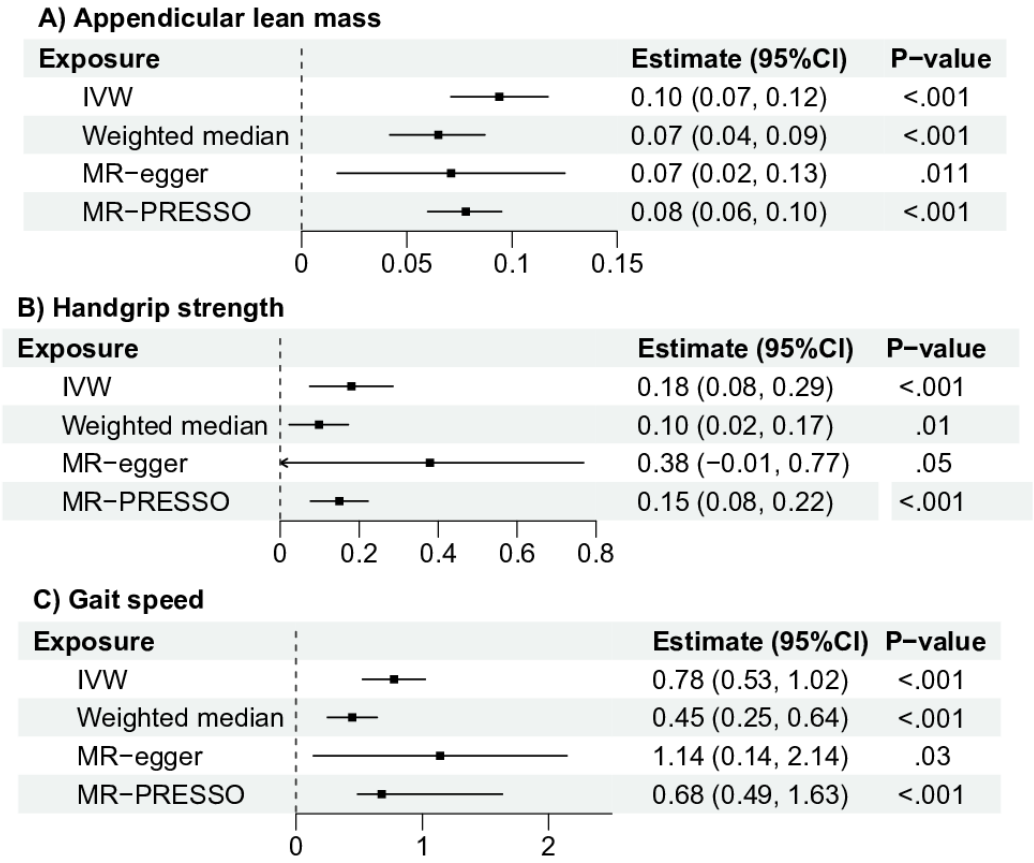
[e]Model 3 was additionally adjusted for smoking, alcohol intake, sleep duration, and TV viewing.

MR Analyses

The 2-Sample MR Analyses

As shown in Figure 2, for an increase of each one unit in genetically predicted ALM, handgrip strength, and gait speed, the general cognitive function score was increased by 0.10 (95% CI 0.07-0.12; *P*<.001), 0.18 (95% CI 0.08-0.29; *P*<.001), and 0.78 (95% CI 0.53-1.02; *P*<.001), respectively. To assess the consistency of directional causation, the effect estimates of 3 different sensitivity methods, that is, weighted median, MR-Egger, and MR-PRESSO were examined and plotted, confirming the directional causation between sarcopenic indices to general cognitive function (see Figure 2 and Figure S2 in Multimedia Appendix 1). There was some evidence of heterogeneity in the SNP effects although the MR-Egger intercepts indicated limited evidence of directional pleiotropy (see Table S3 in Multimedia Appendix 1). In radialMR analyses, outliers were detected (see Figure S3 in Multimedia Appendix 1). MR results remained consistent, showing slightly smaller effects after removing outliers (see Table S4 in Multimedia Appendix 1). For all sarcopenic indices and cognitive function phenotypes, the type 1 error rate was robustly controlled below 0.05 and bias estimates were confined to a narrow range of −0.01 to 0.01 (see Table S5 in Multimedia Appendix 1). Validation through the MRlap method further confirmed that sample overlap did not substantially influence the causal inferences (see Table S6 in Multimedia Appendix 1).

Figure 2. Mendelian randomization results for the relationship of sarcopenic indices with cognitive function. IVW: inverse variance weighted analysis; MR: Mendelian randomization.



Mediation MR Analyses

Figure 3 illustrates the effects of sarcopenic indices on 11 potential mediators. Univariable IVW MR analysis revealed that genetically predicted ALM exhibited protective effects on MVPA (β=0.07, 95% CI 0.05-0.09; *P*<.001) and negative associations with omega-3 fatty acids (β=−0.08, 95% CI −0.12 to −0.05; *P*<.001), falls (β=−0.03, 95% CI −0.05 to -0.003; *P*=.03), and frailty (β=−0.05, 95% CI −0.07 to −0.03; *P*<.001). Genetically predicted handgrip strength showed a positive association with MVPA (β=0.10, 95% CI 0.02-0.19; *P*=.02) and negative associations with falls (β=−0.15, 95% CI −0.24 to −0.07; *P*<.001), and frailty (β=−0.22, 95% CI −0.30 to −0.14; *P*<.001), depression (β=−0.09, 95% CI −0.16 to −0.01; *P*=.02), and stroke (β=−0.41, 95% CI −0.69 to −0.14; *P*=.003). Genetically predicted gait speed was positively associated with MVPA (β=0.83, 95% CI 0.65-1.01; *P*<.001). It also demonstrated significant protective effects against sleep disorders, falls, frailty, anxiety, depression, metabolic syndrome, and type 2 diabetes (with βs ranging from −0.01 to −2.32, all *P* values <.05). Further mediation analysis revealed that the total effect of genetically predicted ALM on general cognitive function decreased from 0.10 (95% CI 0.07-0.12) to 0.09 (95% CI 0.06-0.11) after adjusting for MVPA in MVMR analysis (Table 3). Similarly, the total effect of genetically predicted handgrip strength on general cognitive function attenuated from 0.18 (95% CI 0.08-0.29) to 0.16 (95% CI 0.05-0.27) with adjustment in MVMR analysis. MVPA mediated 8.2% of the total direct effect of ALM on cognitive function, and 10.6% of the total direct effect of handgrip strength on cognitive function. No apparent mediation effect was observed through omega-3 fatty acids, vitamin D level, falls, frailty, sleep disorders, anxiety, depression, stroke, metabolic syndrome, or type 2 diabetes. The type 1 error rate due to sample overlap between sarcopenic indices and the mediators remained below 0.05 (with bias estimates under 0.01) for all phenotypes (see Table S5 in Multimedia Appendix 1).

Figure 3. Effects of genetically predicted sarcopenic indices on potential mediators. MVPA: moderate-to-vigorous intensity physical activity.

Table . The mediation effect of sarcopenic indices on cognitive function via potential mediator in Mendelian randomization (MR) analyses.

Exposures	Mediator	Total effects, β (95% CI)	Direct effects, β (95% CI)	Indirect effects, β (95% CI)	Mediated proportion, % (95% CI)
Appendicular lean mass	MVPA[a]	0.10 (0.07-0.12)[b]	0.09 (0.06-0.11)[b]	0.01 (0.00-0.02)[b]	8.2 (0-16.7)
Handgrip strength	MVPA[a]	0.18 (0.08-0.29)[b]	0.16 (0.05-0.27)[b]	0.02 (0.00-0.05)[b]	10.6 (0-29.6)

[a]MVPA: moderate-to-vigorous intensity physical activity during leisure time.

[b]$P<.05$.

Complementary Analyses

We conducted bidirectional MR analysis to investigate the potential reverse causality between sarcopenia and cognitive function. Table S7 in Multimedia Appendix 1 shows evidence supporting a causal effect of genetically predicted cognitive function on all sarcopenic indices. The effect estimates demonstrated a general consistency across different sensitivity methods (see Figure S2 in Multimedia Appendix 1). Some heterogeneity in the SNP effects was observed, although the MR-Egger intercepts suggested no evidence of directional pleiotropy (see Table S7 in Multimedia Appendix 1). In radialMR analyses, outliers were detected (see Figure S4 in Multimedia Appendix 1). The MR results remained consistent both before and after outlier correction and are presented in Tables S7 and S8 in Multimedia Appendix 1.

Discussion

Principal Findings

This study revealed that sarcopenia and its defining indices (appendicular lean mass, handgrip strength, and gait speed) are associated with cognitive function based on observational data. MR analyses further established a causal relationship between higher levels of sarcopenic indices and better general cognitive function. In addition, physical activity was identified as a significant mediator in the causal pathway linking sarcopenic indices to cognitive function. Our findings suggest sarcopenia as a risk factor and potential biomarker for cognitive impairment, with physical activity offering a therapeutic approach to delay or prevent cognitive decline.

Comparison With Previous Work

The association between sarcopenia and the risk of cognitive impairment has been investigated by several longitudinal studies, but the results were conflicting. Some authors indicated that individuals with sarcopenia [12-17], reduced muscle strength [13,17,35], decreased muscle mass [35], or compromised physical performance [35,36] were associated with an elevated risk of cognitive impairment. However, some others reported no significant association between sarcopenia or its indicators and the risk of cognitive impairment [19-21]. The conflicting results could stem from differences in how sarcopenia is defined and the variations in sample sizes. Our observational analyses addressed these discrepancies by focusing on the more recent EWGSOP2 definition of sarcopenia and its 3 defining indices within a substantial sample size (>200,000 participants). As a result, we found that sarcopenia, as a construct, and its 3 distinct indices were closely correlated with cognitive impairment. Furthermore, the discrepancies observed in previous observational studies might also be attributed to factors such as residual confounding and measurement errors. To mitigate these limitations in our study, we used MR, a genetic epidemiological technique using genetic variants as proxies for the exposure [37]. This approach is less vulnerable to the aforementioned limitations since the genetic variants are accurately measured and documented and are randomly allocated during gamete formation and conception. Therefore, this method can minimize the potential for measurement errors and reduce the likelihood of being influenced by confounding variables [37].

Possible Explanations

Our study demonstrates that physical inactivity is a potential mediating factor in the causal pathway between sarcopenia and cognitive impairment. On one hand, individuals with sarcopenia might have a low level of physical activity [38], which could be attributed to the fact that weakened muscles might hinder the ability to exercise regularly. On the other hand, physical activity positively impacts brain health [39]. First, it can stimulate the formation of new neurons, enhance neuronal survival [10], increase resistance to brain injuries, and facilitate synaptic development and plasticity [40]. Second, physical activity promotes better blood vessel formation in the brain, which is associated with increased learning capabilities [10,11]. Third, it also activates specific gene expression profiles benefiting brain plasticity and cognitive function [41]. Fourth, engaging in physical activity has been associated with reduced amyloid deposition in the brains of cognitively normal elderly adults [42]. Furthermore, physical activity can reduce systemic inflammatory markers [43] and boost the production of neuroprotective proteins like brain-derived neurotrophic factor (BDNF), supporting the growth and survival of neurons [44]. In addition, it positively influences energy balance and glucose metabolism by modulating AMP kinase and insulin signaling, potentially aiding Aβ clearance [45]. These multiple benefits of physical activity contribute to safeguarding cognitive function and underscore the importance of maintaining an active lifestyle in individuals with sarcopenia to support brain health.

Limitations

The triangulation of findings through complementary cohort and MR approaches significantly bolstered the confidence in our drawn inferences. However, several limitations warrant consideration. First, despite using multiple MR methods to resist pleiotropy-related confounding, we could not eliminate residual confounding, which is a known limitation of the MR approach. Second, due to a mere 5% response rate and healthy volunteer bias in the UK Biobank, whether our findings can be generalized

to the broader UK population remains uncertain, despite the large sample size. Third, although we have accounted for common modifiable lifestyle factors and preventable diseases to inform public health policies, this study does not encompass all potential mediation pathways. Fourth, only self-reported gait speed was collected, lacking objective measurements for more accurate estimates. Fifth, the inability to perform stratified analyses by age and gender due to data constraints limits insights into these factors' roles in the sarcopenia-cognition relationship. Sixth, the overlap between GWAS datasets in the MR analysis could bias results and inflate Type 1 error rates. To address this, we assessed the error rate and applied the MRlap method, confirming the robustness of our findings and minimizing the influence of sample overlap on causal associations. Finally, to maintain homogeneity, we focused on individuals of European ancestry in our sample selection, which might influence the generalization of the results, although the sample size was substantial.

Clinical and Research Implications

Our findings provide a promising approach to alleviate the burden of cognitive impairment by identifying and intervening in sarcopenia. Specifically, we have recognized sarcopenia as a risk factor for future cognitive impairment, making it a potential clinical biomarker to screen adults at risk of late-life cognitive impairment. Although no pharmaceutical treatment has been specifically approved for sarcopenia, implementing non-pharmacological interventions, such as physical activity, can serve as a therapeutic approach to proactively delay or prevent the onset of cognitive impairment in affected individuals. Our research indicates that physical activity can mediate the effect of sarcopenia on cognitive function, offering valuable insights that complement the prevailing emphasis on intellectual pursuits as the primary means of exercising the brain [46]. Promoting physical activity may yield a dual positive impact, addressing both sarcopenia and cognitive impairment simultaneously. This approach can potentially enhance the overall health and well-being of those affected by sarcopenia.

Conclusions

Our study suggests that sarcopenia is a causal risk factor for cognitive impairment. Physical activity, a modifiable factor, is capable of measuring the effect of sarcopenia on cognitive function.

Acknowledgments

This work was supported by the National Key Research and Development Program (2022YFC3601900 and 2022YFC2505500), the National Natural Science Foundation of China (81930071, 81902265, 82072502, 82302771, and 82402882), the National Natural Science Foundation Regional Innovation and Development Joint Fund (U21A20352), the Natural Science Foundation of Hunan Province (2022JJ40821, 2024JJ6667), the Project Program of National Clinical Research Center for Geriatric Disorders (Xiangya Hospital; 2020LNJJ03 and 2021LNJJ06), the Scientific Research Program of the Education Department of Hunan Province (296840), and the Natural Science Foundation of Changsha (kq2403017). The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Data Availability

The primary data from the UK Biobank resource are accessible upon application from the UK Biobank repository. The genome-wide association study (GWAS) summary statistics for sarcopenic indices, cognitive function, and potential mediators are publicly available and can be directly accessed through the original publications cited in the reference list [24,26,47-59].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods and results for the study.
[DOCX File, 1104 KB - aging_v8i1e66031_app1.docx]

References

1. Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. JAMA 2014 Dec 17;312(23):2551-2561. [doi: 10.1001/jama.2014.13806] [Medline: 25514304]
2. Pan CW, Wang X, Ma Q, Sun HP, Xu Y, Wang P. Cognitive dysfunction and health-related quality of life among older Chinese. Sci Rep 2015 Nov 25;5:17301. [doi: 10.1038/srep17301] [Medline: 26601612]
3. Fogg C, Griffiths P, Meredith P, Bridges J. Hospital outcomes of older people with cognitive impairment: an integrative review. Int J Geriatr Psychiatry 2018 Jun 26;33(9):1177-1197. [doi: 10.1002/gps.4919] [Medline: 29947150]
4. Rait G, Walters K, Bottomley C, Petersen I, Iliffe S, Nazareth I. Survival of people with clinical diagnosis of dementia in primary care: cohort study. BMJ 2010 Aug 5;341:c3584. [doi: 10.1136/bmj.c3584] [Medline: 20688840]
5. Cooper C, Li R, Lyketsos C, Livingston G. Treatment for mild cognitive impairment: systematic review. Br J Psychiatry 2013 Sep;203(3):255-264. [doi: 10.1192/bjp.bp.113.127811] [Medline: 24085737]

6. Sakal C, Li T, Li J, Li X. Identifying predictive risk factors for future cognitive impairment among Chinese older adults: longitudinal prediction study. JMIR Aging 2024 Mar 22;7:e53240. [doi: 10.2196/53240] [Medline: 38534042]
7. Scarmeas N, Anastasiou CA, Yannakoulia M. Nutrition and prevention of cognitive impairment. Lancet Neurol 2018 Nov;17(11):1006-1015. [doi: 10.1016/S1474-4422(18)30338-7] [Medline: 30244829]
8. Petermann-Rocha F, Balntzi V, Gray SR, et al. Global prevalence of sarcopenia and severe sarcopenia: a systematic review and meta‐analysis. J Cachexia Sarcopenia Muscle 2022 Feb;13(1):86-99. [doi: 10.1002/jcsm.12783] [Medline: 34816624]
9. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European Working Group on sarcopenia in older people. Age Ageing 2010 Jul;39(4):412-423. [doi: 10.1093/ageing/afq034] [Medline: 20392703]
10. van Praag H, Christie BR, Sejnowski TJ, Gage FH. Running enhances neurogenesis, learning, and long-term potentiation in mice. Proc Natl Acad Sci U S A 1999 Nov 9;96(23):13427-13431. [doi: 10.1073/pnas.96.23.13427] [Medline: 10557337]
11. Black JE, Isaacs KR, Anderson BJ, Alcantara AA, Greenough WT. Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. Proc Natl Acad Sci U S A 1990 Jul;87(14):5568-5572. [doi: 10.1073/pnas.87.14.5568] [Medline: 1695380]
12. Lin A, Wang T, Li C, et al. Association of sarcopenia with cognitive function and dementia risk score: a national prospective cohort study. Metabolites 2023 Feb 8;13(2):245. [doi: 10.3390/metabo13020245] [Medline: 36837864]
13. Peng TC, Chiou JM, Chen TF, Chen YC, Chen JH. Grip strength and sarcopenia predict 2-year cognitive impairment in community-dwelling older adults. J Am Med Dir Assoc 2023 Mar;24(3):292-298. [doi: 10.1016/j.jamda.2022.10.015] [Medline: 36435272]
14. Salinas-Rodríguez A, Palazuelos-González R, Rivera-Almaraz A, Manrique-Espinoza B. Longitudinal association of sarcopenia and mild cognitive impairment among older Mexican adults. J Cachexia Sarcopenia Muscle 2021 Dec;12(6):1848-1859. [doi: 10.1002/jcsm.12787] [Medline: 34535964]
15. Hu Y, Peng W, Ren R, Wang Y, Wang G. Sarcopenia and mild cognitive impairment among elderly adults: the first longitudinal evidence from CHARLS. J Cachexia Sarcopenia Muscle 2022 Dec;13(6):2944-2952. [doi: 10.1002/jcsm.13081] [Medline: 36058563]
16. Nishiguchi S, Yamada M, Shirooka H, et al. Sarcopenia as a risk factor for cognitive deterioration in community-dwelling older adults: a 1-year prospective study. J Am Med Dir Assoc 2016 Apr 1;17(4):372. [doi: 10.1016/j.jamda.2015.12.096] [Medline: 26897591]
17. Beeri MS, Leugrans SE, Delbono O, Bennett DA, Buchman AS. Sarcopenia is associated with incident Alzheimer's dementia, mild cognitive impairment, and cognitive decline. J Am Geriatr Soc 2021 Jul;69(7):1826-1835. [doi: 10.1111/jgs.17206] [Medline: 33954985]
18. Amini N, Ibn Hach M, Lapauw L, et al. Meta-analysis on the interrelationship between sarcopenia and mild cognitive impairment, Alzheimer's disease and other forms of dementia. J Cachexia Sarcopenia Muscle 2024 Aug;15(4):1240-1253. [doi: 10.1002/jcsm.13485] [Medline: 38715252]
19. Abellan van Kan G, Cesari M, Gillette-Guyonnet S, et al. Sarcopenia and cognitive impairment in elderly women: results from the EPIDOS cohort. Age Ageing 2013 Mar;42(2):196-202. [doi: 10.1093/ageing/afs173] [Medline: 23221099]
20. van Kan GA, Cesari M, Gillette-Guyonnet S, Dupuy C, Vellas B, Rolland Y. Association of a 7-year percent change in fat mass and muscle mass with subsequent cognitive dysfunction: the EPIDOS-Toulouse cohort. J Cachexia Sarcopenia Muscle 2013 Sep;4(3):225-229. [doi: 10.1007/s13539-013-0112-z] [Medline: 23888380]
21. Taekema DG, Ling CHY, Kurrle SE, et al. Temporal relationship between handgrip strength and cognitive performance in oldest old people. Age Ageing 2012 Jul;41(4):506-512. [doi: 10.1093/ageing/afs013] [Medline: 22374646]
22. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ 2021 Oct 26;375:n2233. [doi: 10.1136/bmj.n2233] [Medline: 34702754]
23. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jul 1;48(4):601-601. [doi: 10.1093/ageing/afz046]
24. Pei YF, Liu YZ, Yang XL, et al. The genetic architecture of appendicular lean mass characterized by association analysis in the UK Biobank study. Commun Biol 2020 Oct 23;3(1):608. [doi: 10.1038/s42003-020-01334-0] [Medline: 33097823]
25. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018 May 30;7:e34408. [doi: 10.7554/eLife.34408] [Medline: 29846171]
26. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 2018 Jul 23;50(8):1112-1121. [doi: 10.1038/s41588-018-0147-3] [Medline: 30038396]
27. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from mendelian randomization analyses with multiple genetic variants. Epidemiology (Sunnyvale) 2017 Jan;28(1):30-42. [doi: 10.1097/EDE.0000000000000559] [Medline: 27749700]
28. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018 May;50(5):693-698. [doi: 10.1038/s41588-018-0099-7] [Medline: 29686387]

29. Bowden J, Spiller W, Del Greco M F, et al. Improving the visualization, interpretation and analysis of two-sample summary data mendelian randomization via the radial plot and radial regression. Int J Epidemiol 2018 Aug 1;47(4):1264-1278. [doi: 10.1093/ije/dyy101] [Medline: 29961852]
30. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample mendelian randomization. Genet Epidemiol 2016 Nov;40(7):597-608. [doi: 10.1002/gepi.21998] [Medline: 27625185]
31. Mounier N, Kutalik Z. Bias correction for inverse variance weighting mendelian randomization. Genet Epidemiol 2023 Jun;47(4):314-331. [doi: 10.1002/gepi.22522] [Medline: 37036286]
32. MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002 Mar;7(1):83-104. [doi: 10.1037/1082-989x.7.1.83] [Medline: 11928892]
33. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods 2013 Jun;18(2):137-150. [doi: 10.1037/a0031034] [Medline: 23379553]
34. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017 Nov;13(11):e1007081. [doi: 10.1371/journal.pgen.1007081] [Medline: 29149188]
35. Amini N, Dupont J, Lapauw L, et al. Sarcopenia-defining parameters, but not sarcopenia, are associated with cognitive domains in middle-aged and older European men. J Cachexia Sarcopenia Muscle 2023 Jun;14(3):1520-1532. [doi: 10.1002/jcsm.13229] [Medline: 37021434]
36. Moon JH, Moon JH, Kim KM, et al. Sarcopenia as a predictor of future cognitive impairment in older adults. J Nutr Health Aging 2016;20(5):496-502. [doi: 10.1007/s12603-015-0613-x] [Medline: 27102786]
37. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008 Apr 15;27(8):1133-1163. [doi: 10.1002/sim.3034]
38. Chiba I, Lee S, Bae S, et al. Difference in sarcopenia characteristics associated with physical activity and disability incidences in older adults. J Cachexia Sarcopenia Muscle 2021 Dec;12(6):1983-1994. [doi: 10.1002/jcsm.12801] [Medline: 34612020]
39. Mc Ardle R, Taylor L, Cavadino A, Rochester L, Del Din S, Kerse N. Characterizing walking behaviors in aged residential care using accelerometry, with comparison across care levels, cognitive status, and physical function: cross-sectional study. JMIR Aging 2024 Jun 4;7:e53020. [doi: 10.2196/53020] [Medline: 38842168]
40. Lu B, Chow A. Neurotrophins and hippocampal synaptic transmission and plasticity. J Neurosci Res 1999 Oct 1;58(1):76-87. [Medline: 10491573]
41. Escorihuela RM, Tobeña A, Fernández-Teruel A. Environmental enrichment and postnatal handling prevent spatial learning deficits in aged hypoemotional (Roman high-avoidance) and hyperemotional (Roman low-avoidance) rats. Learn Mem 1995;2(1):40-48. [doi: 10.1101/lm.2.1.40] [Medline: 10467565]
42. Brown BM, Peiffer JJ, Martins RN. Multiple effects of physical activity on molecular and cognitive signs of brain aging: can exercise slow neurodegeneration and delay Alzheimer's disease? Mol Psychiatry 2013 Aug;18(8):864-874. [doi: 10.1038/mp.2012.162] [Medline: 23164816]
43. Ford ES. Does exercise reduce inflammation? Physical activity and C-reactive protein among U.S. adults. Epidemiology 2002 Sep;13(5):561-568. [doi: 10.1097/00001648-200209000-00012] [Medline: 12192226]
44. Nichol K, Deeny SP, Seif J, Camaclang K, Cotman CW. Exercise improves cognition and hippocampal plasticity in APOE epsilon4 mice. Alzheimers Dement 2009 Jul;5(4):287-294. [doi: 10.1016/j.jalz.2009.02.006] [Medline: 19560099]
45. De Felice FG, Vieira MNN, Bomfim TR, et al. Protection of synapses against Alzheimer's-linked toxins: insulin signaling prevents the pathogenic binding of abeta oligomers. Proc Natl Acad Sci U S A 2009 Feb 10;106(6):1971-1976. [doi: 10.1073/pnas.0809158106] [Medline: 19188609]
46. Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. Nat Rev Neurosci 2008 Jan;9(1):58-65. [doi: 10.1038/nrn2298] [Medline: 18094706]
47. Hand grip strength (right). IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-10215/ [accessed 2025-05-21]
48. Usual walking pace. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-4711/ [accessed 2025-05-21]
49. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2025-05-21]
50. Wang Z, Emmerich A, Pillon NJ, et al. Genome-wide association analyses of physical activity and sedentary behavior provide insights into underlying mechanisms and roles in disease prevention. Nat Genet 2022 Sep;54(9):1332-1344. [doi: 10.1038/s41588-022-01165-1] [Medline: 36071172]
51. Non-cancer illness code, self-reported: anxiety/panic attacks. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/ukb-b-17243/ [accessed 2015-05-21]
52. Trajanoska K, Seppala LJ, Medina-Gomez C, et al. Genetic basis of falling risk susceptibility in the UK Biobank Study. Commun Biol 2020 Sep 30;3(1):543. [doi: 10.1038/s42003-020-01256-x] [Medline: 32999390]
53. Atkins JL, Jylhävä J, Pedersen NL, et al. A genome-wide association study of the frailty index highlights brain pathways in ageing. Aging Cell 2021 Sep;20(9):e13459. [doi: 10.1111/acel.13459] [Medline: 34431594]
54. Jiang X, O'Reilly PF, Aschard H, et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat Commun 2018 Jan 17;9(1):260. [doi: 10.1038/s41467-017-02662-2] [Medline: 29343764]

55. Study: GCST009602. GWAS catalog. URL: https://www.ebi.ac.uk/gwas/studies/GCST009602 [accessed 2025-05-21]
56. Omega-3 fatty acids. IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/met-d-Omega_3/ [accessed 2025-05-21]
57. Sleep disorders (combined). IEU open GWAS project. URL: https://gwas.mrcieu.ac.uk/datasets/finn-b-SLEEP/ [accessed 2025-05-21]
58. Malik R, Traylor M, Pulit SL, et al. Low-frequency and common genetic variation in ischemic stroke: the METASTROKE collaboration. Neurology (ECronicon) 2016 Mar 29;86(13):1217-1226. [doi: 10.1212/WNL.0000000000002528] [Medline: 26935894]
59. Mahajan A, Wessel J, Willems SM, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet 2018 Apr;50(4):559-571. [doi: 10.1038/s41588-018-0084-1] [Medline: 29632382]

Abbreviations

A level: Advanced Level
ALM: appendicular lean mass
AS level: Advanced Subsidiary Level
EWGSOP2: European Working Group on Sarcopenia in Older People in 2019
GCSE: General Certificate of Secondary Education
GWAS: genome-wide association study
HNC: Higher National Certificate
HND: Higher National Diploma
IVW: inverse variance weighted analysis
MR: Mendelian randomization
MVMR: multivariable Mendelian randomization
MVPA: moderate-to-vigorous intensity physical activity
NVQ: National Vocational Qualification
OR: odds ratio
SNP: single nucleotide polymorphism

Edited by R Bjarnadottir; submitted 02.09.24; peer-reviewed by G Feng, Z Wang; revised version received 12.03.25; accepted 21.04.25; published 11.06.25.

Please cite as:
Sha T, Zhang Y, Wei J, Li C, Zeng C, Lei G, Wang Y
Sarcopenia and Risk of Cognitive Impairment: Cohort Study and Mendelian Randomization Analysis
JMIR Aging 2025;8:e66031
URL: https://aging.jmir.org/2025/1/e66031
doi:10.2196/66031

Feasibility of a Cinematic–Virtual Reality Program Educating Health Professional Students About the Complexity of Geriatric Care: Pilot Pre-Post Study

Elizabeth A Beverly[1,2*], PhD; Samuel Miller[3*], BA; Matthew Love[2*], MFA; Carrie Love[2], MFA

[1]Department of Primary Care, Ohio University Heritage College of Osteopathic Medicine, 1 Ohio University, Athens, OH, United States

[2]The Diabetes Institute, Ohio University, Athens, OH, United States

[3]Department of Medicine, Ohio University Heritage College of Osteopathic Medicine, Athens, OH, United States

[*]these authors contributed equally

Corresponding Author:
Elizabeth A Beverly, PhD
Department of Primary Care, Ohio University Heritage College of Osteopathic Medicine, 1 Ohio University, Athens, OH, United States

Abstract

Background: The US population is aging. With this demographic shift, more older adults will be living with chronic conditions and geriatric syndromes. To prepare the next generation of health care professionals for this aging population, we need to provide training that captures the complexity of geriatric care.

Objective: This pilot study aimed to assess the feasibility of the cinematic–virtual reality (cine-VR) training in the complexity of geriatric care. We measured changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy before and after participating in the training program.

Methods: We conducted a single-arm, pretest-posttest pilot study to assess the feasibility of a cine-VR training and measure changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. Health professional students from a large university in the Midwest were invited to participate in 1 of 4 cine-VR trainings. Participants completed 3 surveys before and after the cine-VR training. We performed paired *t* tests to examine changes in these constructs before and after the training.

Results: A total of 65 health professional students participated in and completed the full cine-VR training for 100% retention. Participants did not report any technological difficulties or adverse effects from wearing the head-mounted displays or viewing the 360-degree video. Out of the 65 participants, 48 completed the pre- and postassessments. We observed an increase in awareness of discrimination towards people with disability (t_{47}=−3.97; P<.001). In addition, we observed significant improvements in self-efficacy to identify and manage elder abuse and neglect (t_{47}=−3.36; P=.002). Finally, we observed an increase in participants' empathy (t_{47}=−2.33; P=.02).

Conclusions: We demonstrated that our cine-VR training program was feasible and acceptable to health professional students at our Midwestern university. Findings suggest that the cine-VR training increased awareness of discrimination towards people with disabilities, improved self-efficacy to identify and manage elder abuse and neglect, and increased empathy. Future research using a randomized controlled trial design with a larger, more diverse sample and a proper control condition is needed to confirm the effectiveness of our cine-VR training.

(JMIR Aging 2025;8:e64633) doi:10.2196/64633

KEYWORDS

virtual reality; VR; aging; geriatric syndromes; diabetes; elder abuse and neglect; gerontology; geriatrics; older; elderly; education; student; cinematic; video; head mounted; feasibility; experience; attitude; opinion; perception; elder abuse; chronic conditions; older adult care; health intervention; randomized controlled trial

Introduction

The US population is aging. Over the next 3 decades, the number of Americans aged 65 and older is projected to increase from 58 million in 2022 to 82 million in 2050, representing a 47% increase [1]. With this demographic shift, more older adults will be living with chronic conditions. According to the National Council on Aging, 94.9% of adults aged 60 years and older have at least one chronic condition and 78.7% have two or more [2]. Furthermore, 42% of adults aged 60 years and older have

obesity, which increases the risk for cardiovascular disease, type 2 diabetes, and certain cancers [3]. This translates to older Americans requiring more health care and access to health care professionals. Thus, training the next generation of health care professionals in the complexity of care for older adults is an urgent need.

Despite this need, the education system lacks adequate training in geriatric care [4]. Health professional training should address age-related physiological, physical, cognitive, and psychosocial changes that may negatively impact quality of life, disability, morbidity, and mortality [5,6]. Furthermore, this training should emphasize the early detection of and treatment for geriatric syndromes (eg, polypharmacy, falls, frailty, and incontinence) given their high prevalence and frequent under-recognition in older adults [7]. Finally, health professional training should draw attention to implicit and explicit biases in aging as well as ethical issues related to elder abuse and neglect and end-of-life care [8].

Clinical simulation may be an innovative approach to teaching health professional students about the complexity of care in older adults. Previous research demonstrates that clinical simulation is an effective methodology for health professional students [9]. The value of clinical simulation is that it replaces or enhances real experiences with guided experiences that replicate the real world in a fully interactive manner [10]. Furthermore, clinical simulation is repeatable, which allows for the repetition of acquired skills with appropriate feedback [11]. Clinical simulation can be delivered in different modalities, including role play, the use of standardized patients, computer screen-based simulation, and the use of electronic patient records [11]. Electronic patients are most commonly mannequin-based or replicas of clinical sites; however, new advancements in technology have led to more sophisticated and realistic virtual reality (VR) based simulations [12].

Cinematic-VR (cine-VR) is one type of VR-based simulation. Specifically, cine-VR leverages 360-degree video and spatialized audio through film. It differs from traditional VR, which uses computer-generated characters and worlds. In cine-VR, filmmakers apply the techniques of cinema (ie, narrative storytelling, script writing, actors, lighting, and postproduction) to create an immersive, visually compelling environment [13,14]. We created a cine-VR training program to depict the complexity of geriatric care. Our clinical simulation depicted an older adult with multimorbidity, disability, and several geriatric syndromes. In addition, we wove in themes of intersectionality with regards to ageism, ableism, and racism. We conducted a pilot study to assess the feasibility and acceptability of the cine-VR training program with health professional students. We also measured changes in health professional students' attitudes toward disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. The pilot study was designed to assess the feasibility of our data collection methods, willingness of participants to complete the cine-VR training program, sample size estimation, and refinement of measurements. We will use the findings to inform the development of a future randomized controlled trial to compare the cine-VR training program to traditional instruction.

Methods

Research Design

We conducted a single-arm, pretest-posttest study to assess changes in attitudes to disability, self-efficacy to identify and manage elder abuse and neglect, and empathy.

Ethical Considerations

We obtained ethics approval from the Ohio University Office of Research Compliance Institutional Review Board (approval number: #22-X-153). We ensured our pilot study met the requirements set forth in the regulations on public welfare in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) by complying with federal, state, and local laws and regulations for human subjects; the principles set forth in "The Belmont Report," and the Helsinki Declaration of 1975. All participants provided electronic informed consent. Informed consent was completed before participation in the study.

Cine-VR Episodes

The cine-VR episodes were designed to educate health professional students about complex health conditions, social drivers of health, and implicit bias. We conveyed these objectives through our patient, John Chen, an 80-year-old Chinese American man with a 14-year history of type 2 diabetes and comorbid hypertension. Mr. Chen is a person who is hard of hearing with a mobility disability. His most recent HBA_{1c} level was 9.7%. His medications include metformin and glyburide for his diabetes and lisinopril for his blood pressure. He also has osteoarthritis, decreased renal function (eGFR=58 mL/min/1.73 m^2), elevated triglycerides due to his glucose levels (299 mg/dL), and a body mass index of 23.1 kg/m^2. He has not received diabetes self-management education and support, despite being enrolled in both Medicare and Medicaid. Recently, Mr Chen moved in with his son's family, but he is having a difficult time adjusting to his new home. Although he is able to move around on his own, his ambulation is unsteady, and he falls often. He also experiences frequent urinary incontinence, which frustrates his son and daughter-in-law because they have to clean him. Furthermore, Mr Chen does not drive, shop, cook, or bathe on his own. Therefore, he is reliant on his family to meet most of his daily needs. This leaves Mr Chen feeling isolated, dependent, and ashamed of his physical limitations.

In the cine-VR training program, participants watched 6 episodes, ranging in length from 2 minutes to 5 minutes. The episodes captured 3 separate patient-provider interactions, including visits with a primary care physician, an urgent care physician, and an emergency department physician. Participants were encouraged to identify issues with his medical care as well as concerns about the caregiving he received at home. Specifically, participants were challenged to figure out why Mr Chen was experiencing so many falls, injuries, and health emergencies.

The fifth and sixth episodes of the cine-VR training program were "guided simulations," or prerecorded cine-VR face-to-face conversations with the emergency department physician and Mr Chen. The guided simulations were designed to simulate a

high-stakes conversation, to give participants an opportunity to practice difficult conversations without the pressure of causing harm. Participants were instructed to speak predetermined dialogue that appeared at the bottom of the 360-degree video visible within the headset. The first "guided simulation" was a conversation between the participant and the emergency department physician. The participant assumed the role of a clinical colleague and encouraged the physician to inquire further about the bruises observed on John's body and to call social work for suspicion of elder abuse and neglect. The second "guided simulation" was a conversation between the emergency department physician, Mr Chen, and a social worker. In this "guided simulation," the participant took on the role of the emergency department physician and asked Mr Chen difficult questions about his injuries, his family, and life at home.

Cine-VR Training Curricular Content

We developed curricular content to be taught in tandem with the cine-VR episodes. The curriculum included 6 debriefs or reflections that reinforced key takeaways from each cine-VR episode. The key takeaways from each episode focused on the following content: (1) Diabetes, disability, and the aging population, (2) bias toward disability, (3) association between disability and elder abuse and neglect, (4) recognizing elder abuse and neglect, (5) identifying risk factors for elder abuse and neglect, and (6) reporting elder abuse and neglect. An experienced behavioral diabetes researcher (EAB) delivered both trainings. The integrity of the curricular content was ensured by written materials, a peer-review process of all materials, delivery by one trained behavioral diabetes expert, and team member observation of cine-VR training.

Cine-VR Technology

Participants viewed the cine-VR episodes with Pico G2 4K head-mounted displays. These head-mounted displays allowed participants to move their head and body in any direction to choose what aspects they paid attention to during each cine-VR episode. We also synchronized all cine-VR episodes from a central computer using VR Sync software so all participants viewed the same content in the head-mounted display at the same time.

Power Analysis

This was a pilot study so we did not conduct an a priori power analysis. Using the guidance of Lancaster et al [15], we recruited a minimum sample size of 30 participants.

Recruitment

Health professional students were recruited from a large university in the Midwest. Eligibility criteria for participating in the pilot study included English-speaking and reading adults aged 18 years and older who were enrolled during the 2022 - 2023 academic year and majoring in one of the following programs: premedicine program, speech-language pathology, audiology, nutrition and dietetics, physical therapy, nursing (any level), social work (any level), athletic training, exercise physiology, pre-physician assistant program, psychology, public health administration, or public health. There were no other exclusion criteria. Specifically, we hosted 4 in-person cine-VR training sessions on campus in classrooms and invited students from those health professional programs to participate. Sessions were hosted on October 17, 2022, October 24, 2022, April 5, 2023, and June 6, 2023.

Measures

To assess the feasibility of the cine-VR training program, we measured recruitment, retention, length of time required to recruit, rate of completion of the cine-VR training, and feasibility of the data collection measures. In addition, participants provided demographic factors (age, gender, race, ethnicity, year in program, and program) and completed the following measures:

Attitudes to Disability Scale

The Attitudes to Disability Scale [16] is a 16-item validated scale that assesses attitudes toward disability and attitudes toward people with a disability. The 16 items were answered on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The scale included 4 domains or subscales with good internal consistency: Inclusion—focuses on attitudes toward the inclusion and exclusion for people with disabilities as well as the burden on families and society (Cronbach α=0.76), Discrimination—focuses on awareness of unfair and prejudicial treatment to people with disabilities (Cronbach α=0.74), Gains—focuses on positives of having a disability (Cronbach α=0.75), and Prospects—focuses on future hopes and whether disability impacts those hopes (Cronbach α=0.72) [16].

Responding to Elder Abuse in GERiAtric Care—Provider Questionnaire

The Responding to Elder Abuse in GERiAtric Care—Provider Questionnaire [17] is a validated 25-item scale that assesses health care provider preparedness to identify and manage elder abuse and neglect. For the purposes of this study, we used 8 of the 25 items. The items we used addressed self-efficacy to identify and manage elder abuse and neglect on a 0 to 10 scale. The 17 excluded items were written for practicing health care professionals. In our sample, the 8 self-efficacy questions demonstrated excellent internal consistency for identifying and managing elder abuse and neglect (Cronbach α=0.95) [17].

Jefferson Scale of Empathy Health Professionals Students Version

The Jefferson Scale of Empathy Health Professionals Students Version [18] is a 20-item validated scale that assesses empathy among health professional students. The 20 items were answered on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). Scores are summed for a composite score, ranging from 20 to 140. The scale demonstrated good internal consistency (Cronbach α=0.85) [18].

Data Collection

Participants completed the assessments through Qualtrics, an electronic questionnaire service. To access the preassessment, participants scanned a QR code from a Microsoft PowerPoint slide that directed them to the Qualtrics website. The Qualtrics website included a brief description of the study, the informed consent form, demographic questions, and the 3 measures. All participants provided informed consent without a signature in Qualtrics before completing the measures. After the cine-VR

training, participants scanned a different QR code that directed them to the postassessment measures.

We collected data anonymously. To link participants' pre- and postassessment responses, we included 3 questions at the beginning of the pre- and postassessments, which served as a unique identifier (ie, model of first car, high school mascot, and the day of the month on which they were born [eg, 31]). Total time to complete the pre- and postmeasures took approximately 15 - 20 minutes. Participants received a US $25 gift card for human subject compensation. To maintain anonymity, participants clicked on a new Qualtrics link that was not connected to their pre- or postmeasures to receive a gift card; it was possible that participants received compensation without completing all pre- and postmeasures.

Statistical Analysis

First, we assessed participants' demographic factors using descriptive statistics and presented them as means and SDs or sample size and percentages. Next, we examined the distribution of the data to ensure they met statistical test assumptions. To assess changes pre- and post–cine-VR training, we conducted paired *t* tests; all test statistics are presented with α values and degrees of freedom. We also calculated effect sizes using Cohen *d*, with a small effect=0.2, medium effect=0.5, and a large effect=0.8. We defined statistical significance as a *P* value less than .05 and conducted analyses in IBM SPSS statistical software (version 29.0).

Results

Feasibility

A total of 65 health professional students participated in and completed the full cine-VR training for 100% retention. Participants did not report any technological difficulties or adverse effects from wearing the head-mounted displays or viewing the 360-degree video. We scheduled 4 cine-VR trainings in a 6-month period to recruit the 65 participants, and all 4 trainings lasted approximately 60 minutes. For data collection, 65 participants completed the preassessment measures; however, only 48 completed the postassessment measures, resulting in a 74% completion rate. The 74% rate of completion suggests that we may want to consider adaptations or refinements to our selected measures. Overall, the ease of recruitment, length of time required to recruit, retention in the cine-VR trainings, and feasibility of data collection methods suggest that the cine-VR training was feasible.

Demographics

A total of 48 health professional students consented to participate in the pilot study and completed all pre- and postassessments. The mean age of participants was 22.5 (SD 3) (see Table 1). A total of 41 participants (85.4%) self-identified as women, 6 (12.5%) self-identified as men, and 1 (2.1%) self-identified as nonbinary. The participants self-identified their race as follows: 6.3% (3/48) Asian or Pacific Islander, 18.8% (9/48) Black or African American, 4.2% (2/48) Mixed Race, and 70.8% (34/48) White. For ethnicity, 2 participants (4.2%) self-identified as Hispanic, Latino, or Spanish origin and 2 (4.2%) self-identified as Mexican, Mexican American, or Chicano. The majority of participants were in the final year of their program (37.5%, 18/48), with premedicine as the most reported major (42.6%, 20/48).

Table . Participant demographic characteristics in the cinematic-virtual reality (cine-VR) study (N=48).

Variable	Results
Age (years), mean (SD)	22.5 (3)
Gender, n (%)	
Woman	41 (85.4)
Man	6 (12.5)
Nonbinary	1 (2.1)
Transgender	0 (0)
Genderqueer	0 (0)
An identity not listed	0 (0)
Ethnicity, n (%)	
Hispanic, Latino, or Spanish origin	2 (4.2)
Mexican, Mexican American, or Chicano	2 (4.2)
Non-Hispanic, Non-Latino, or Non-Spanish	44 (91.6)
Race, n (%)	
American Indian or Pacific Islander	0 (0)
Asian or Pacific Islander	3 (6.3)
Black or African American	9 (18.8)
Middle Eastern	0 (0)
Mixed Race	2 (4.2)
White	34 (70.8)
Program [a], n (%)	
Child Life Specialist	18 (38.3)
Exercise Physiology	1 (2.1)
Nutrition	4 (8.5)
Pre-Medicine	20 (42.6)
Pre-Physician Assistant	2 (4.3)
Psychology	2 (4.3)
Year in Program [b], n (%)	
Year 1	9 (18.8)
Year 2	10 (20.8)
Year 3	9 (18.8)
Year 4	18 (37.5)

[a]Missing values “Program” (n=1).

[b]"Year in Program” (n=2).

Attitudes Toward Disability Findings

Pre- and postdomain scores for the Attitudes to Disability Scale can be found in Table 2. Post–cine-VR training, we observed significant changes in 1 of the 4 Attitudes to Disability Subscales: Discrimination Domain (t_{47}=−3.97, P<.001). This change had a Cohen d=0.55, indicating a medium effect. This finding suggests that the cine-VR training may have increased participants’ awareness of discrimination towards people with a disability. In the Discrimination Domain, 3 of the 4 items showed a significant change with more participants “agreeing” and “strongly agreeing” with the following statements posttraining: “People often make fun of disabilities” (t_{47}=−3.92; P<.001); “People tend to become impatient with those with a disability” (t_{47}=−3.19; P=.003); and “People tend to treat those with a disability as if they have no feelings” (t_{47}=−2.69; P=.01). Post-cine-VR training, we did not observe significant changes in the Inclusion Domain (t_{47}=−.16; P=.88), Gains Domain (t_{47}=−1.76; P=.08), or Prospects Domain (t_{47}=−.30; P=.77).

Table . Participants' attitudes to disability scale mean scores before and after the cinematic-virtual reality (cine-VR) training program (N=48).

Questions	Pre-VR[a], mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value
Inclusion domain				
People with a disability find it harder than others to make new friends.	3.9 (0.6)	3.8 (0.8)	.948 (47)	.35
People with a disability have problems getting involved in society.	3.5 (1.2)	3.8 (0.8)	−2.047 (47)	.046
People with a disability are a burden on society.	1.2 (0.4)	1.3 (0.6)	−1.000 (47)	.32
People with a disability are a burden on their family.	1.9 (0.9)	1.7 (1.0)	1.944 (47)	.06
Total	2.6 (0.6)	2.6 (0.5)	−.156 (47)	.88
Discrimination domain				
People often make fun of disabilities.	3.9 (0.6)	4.2 (0.6)	−3.923 (47)	<.001
People tend to become impatient with those with a disability.	3.9 (0.8)	4.3 (0.7)	−3.186 (47)	.003
People tend to treat those with a disability as if they have no feelings.	3.5 (1.1)	3.9 (0.8)	−2.687 (47)	.01
People with a disability are easier to take advantage of (exploit or treat badly) compared with other people.	3.7 (0.9)	3.9 (1.0)	−1.095 (47)	.28
Total	3.8 (0.6)	4.1 (0.6)	−3.967 (47)	<.001
Gains domain				
Having a disability can make someone a stronger person.	3.8 (0.7)	3.9 (0.8)	−1.219 (47)	.23
Having a disability can make someone a wiser person.	3.7 (0.8)	3.7 (0.8)	.227 (47)	.82
Some people achieve more because of their disability (eg, they are more successful).	3.0 (0.8)	3.2 (0.8)	−1.741 (47)	.09
People with a disability are more determined than others to reach their goals.	3.0 (0.7)	3.1 (0.6)	−1.401 (47)	.17
Total	3.3 (0.5)	3.5 (0.6)	−1.764 (47)	.08
Prospects domain				
Sex should not be discussed with people with disabilities.	1.6 (0.9)	1.4 (0.7)	1.752 (47)	.09
People should not expect too much from those with a disability.	1.5 (0.5)	1.6 (0.8)	−1.182 (47)	.24
People with a disability should not be optimistic (hopeful) about their future	1.3 (0.7)	1.5 (0.9)	−1.543 (47)	.13
People with a disability have less to look forward to than others.	1.6 (0.7)	1.6 (0.7)	.423 (47)	.67

Questions	Pre-VR[a], mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value
Total	1.5 (0.5)	1.5 (0.6)	−.300 (47)	.77

[a]VR: virtual reality.

Self-Efficacy to Identify and Manage Elder Abuse and Neglect Findings

Post—cine-VR training, we observed significant changes in participants' self-efficacy to identify and manage elder abuse and neglect (t_{47}=−3.36, *P*=.002, Table 3). This change had a Cohen *d*=0.49, approaching a medium effect.

Table . Participants' perceived self-efficacy for identifying and managing elder abuse and neglect before and after the cinematic-virtual reality (cine-VR) training program (N=48).

Questions	Pre-VR,[a] mean (SD)	Post-VR, mean (SD)	*t* test (*df*)	*P* value	Cohen *d*
Asking questions about abuse to an older patient who has clear indications of now being, or having previously been, subjected to abuse.	6.2 (2.4)	7.2 (2.4)	−3.354 (47)	.002	.48
Asking questions about abuse to an older patient who has no clear indications of now being or having previously been, subjected to abuse.	5.7 (2.6)	5.9 (2.4)	−.585 (47)	.56	.08
Ensuring you are able to ask questions about abuse in private to an older patient who has a relative who insists on being present during all contact.	6.1 (2.7)	7.0 (2.4)	−2.911 (47)	.005	.42
In conversation, providing support to an older patient who tells about abuse.	7.4 (2.6)	8.0 (2.1)	−2.438 (47)	.02	.35
Helping an older patient subjected to abuse on to the right body in health care, or to the right support function in society.	6.6 (2.9)	7.7 (2.1)	−3.174 (47)	.003	.46
Helping an older patient subjected to abuse to make a report to the police or social services.	6.9 (2.8)	7.6 (2.2)	−1.984 (47)	.05	.29
Helping and supporting an older patient subjected to abuse, who does not currently want to change his or her situation.	5.5 (2.9)	6.3 (2.6)	−2.433 (47)	.02	.35
Handling the meeting with an older patient who says no to questions about abuse, but where you still have strong suspicions that the patient is subjected to abuse.	5.7 (2.8)	6.6 (2.2)	−2.601 (47)	.01	.38
Total	6.3 (2.4)	7.0 (1.8)	−3.364 (47)	.002	.49

[a]VR: virtual reality.

In total, 6 of the 8 items demonstrated significant improvements in perceived self-efficacy. Specifically, participants reported increases in their self-efficacy to ask questions about abuse to an older individual who has clear indications of being abused (t_{47}=−3.354; *P*=.002; Cohen *d*=0.48) and to ensure that they ask questions about abuse in private (t_{47}=−2.911; *P*=.005; Cohen *d*=0.42). We also observed increases in self-efficacy for providing support to an older individual who discloses abuse (t_{47}=−2.438; *P*=.02; Cohen *d*=0.35) and directing an older adult experiencing abuse to the right support person or service

(t_{47}=−3.174; P=.003; Cohen d=0.46). Finally, participants showed improvements in their self-efficacy for helping older adults experiencing abuse who do not want to change their situation (t_{47}=−2.433; P=.02; Cohen d=0.35) and handling an older person who says no to questions about abuse despite strong suspicions of abuse (t_{47}=−2.601; P=.01; Cohen d=0.38).

One item showed a trend in improving self-efficacy to help an older adult make a report to the police or social services (t_{47}=−1.984; P=.05; Cohen d=0.29). Participants did not show an improvement in their self-efficacy to ask questions about abuse to an older adult who has no clear indications of presently or previously experiencing abuse (t_{47}=−.585; P=.56; Cohen d=0.08). This finding highlights an area to be addressed in future research.

Empathy Findings

Pre–cine-VR training empathy scores ranged from 47 to 128, with a mean of 105.7 (SD 16.6). Post–cine-VR training empathy scores ranged from 80 to 131, with a mean of 110.4 (SD 12.1). We observed a significant increase in participants' empathy scores post–cine-VR training (mean change=4.8; t_{47}=−2.329; P=.02), with a Cohen d=0.34 indicating a small effect. This finding indicates a noticeable difference in empathy before and after participating in the cine-VR training. While the effect size is small, an increase in empathy may have practical significance in training health professional students about the complexity of diabetes management. More research is needed to confirm this finding.

Discussion

Principal Findings

In this pilot study, we assessed the feasibility of implementing a cine-VR training program with health professional students. Our findings showed that we were able to recruit, implement, and retain participants for the full cine-VR training. Furthermore, participants reported no adverse effects or issues with the cine-VR technology. In total, 3-quarters of the participants completed the pre- and postmeasures for the pilot study. Overall, we observed significant improvements in participants' awareness of discrimination toward people with disability, self-efficacy to identify and manage elder abuse and neglect, and empathy. These improvements suggest that the cine-VR training program may be an effective teaching modality to educate health professional students about the complexity of care in older adults. More research is needed with a larger, more diverse sample and a proper attention control condition to confirm its effectiveness.

An objective of the cine-VR training program was to raise awareness about discrimination towards aging and disabilities. The cine-VR training portrayed multiple interactions among Mr Chen, his family, and health care providers to raise awareness about discrimination. In one simulation, for instance, Mr Chen's grandson is depicted yelling "Good morning!" and laughing, a reaction to his grandfather's hearing loss. Later in the same simulation, his daughter-in-law appears visibly frustrated following an incident of urinary incontinence. These interactions were included to show that sometimes people, including family members, make fun of disabilities or become impatient with people who have a disability. Similarly, the cine-VR training included episodes involving the primary care physician and urgent care physician, directing their questions about Mr Chen's medical history to his son and daughter-in-law, rather than speaking to Mr Chen directly. This was done intentionally to suggest that the providers did not think Mr Chen was capable of discussing his own health. In addition, we included an insensitive comment from the urgent care physician directed at Mr Chen poking fun of his situation. These scenarios were designed to illustrate various types of discrimination and elicit emotional reactions from the participants. Learning that involves emotions increases memory retention through emotional engagement, enhanced cognitive processing, and memory consolidation [19,20]. The brain tends to prioritize emotionally significant information for long-term memory over neutral information, which may explain the increased awareness of discrimination observed in the findings [21].

The cine-VR training program also emphasized recognizing the signs of elder abuse and neglect, how to report abuse, and the subsequent actions taken after a report is made. Each episode uncovered new signs of elder abuse and neglect involving Mr Chen (eg, withholding fluids throughout the day to prevent urinary incontinence, not refilling Mr Chen's medications, and making Mr Chen live in the garage while the house is being renovated), alongside behaviors commonly exhibited by abusers. The follow-up debriefs and reflections concentrated on the responsibilities of mandated reporters, the reporting process, and the steps that follow a report. This content likely contributed to the participants' increased self-efficacy in asking older adults about abuse, ensuring questions asked about abuse are asked in private, providing emotional support to an older adult who tells you about abuse, helping an older adult subjected to abuse to the right authorities, helping an older adult subjected to abuse who does not want to change circumstances, and handling an older adult who denies abuse but you have strong suspicions that they are subjected to abuse. Additional research is needed to confirm if these increases in self-efficacy are sustained over time.

Comparison With Previous Work

Medical training uses a variety of simulation types to prepare health professional students for clinical practice. The most common types of simulation include manikins and advanced patient simulators, standardized patients, and task trainers [22]. The first manikin-based simulation can be traced back to the 1960s with the development of Resusci Anne for CPR training [23]. Since then, health professional students have used manikins and advanced patient simulators to replicate complex medical scenarios, including cardiac arrest, childbirth delivery, and trauma resuscitation [23]. The advantages of using manikins and advanced patient simulators include a risk-free setting for repeated practice of clinical skills and standardized training experiences [24-26]. Drawbacks of using manikins and advanced patient simulators include the high purchase price, ongoing maintenance costs, and the technological limitations that prevent them from fully replicating the complexity and unpredictability of human behavior [27,28]. Conversely, standardized patients

provide a more realistic interaction for students. Standardized patients have been shown to improve communication skills, critical thinking, self-efficacy, and clinical examination skills [29]. Despite training, standardized patients may show individual differences that can lead to variations in role portrayal, impacting the consistency of the simulated learning experience [30]. Unlike manikins, advanced patient simulators, and standardized patients, a task trainer is a specific device designed to help train health professional students on how to perform specific procedures, such as inserting an intravenous line or performing a lumbar puncture. Task trainers allow for focused learning, skill mastery, and standardization of training in a safe learning environment equipped to provide feedback and assessment; however, they are limited in scope, lack human interaction, and require regular maintenance and upkeep to remain functional [22].

Cine-VR builds on these traditional simulation methods by integrating narrative storytelling and interactive visual experiences [13,14,31,32]. Unlike manikins, advanced patient simulators, and task trainers, cine-VR offers an immersive environment that shows real-life scenarios, situational complexity, and human emotions [33]. Furthermore, cine-VR can provide both physical and decision-based feedback through interactive episodes, whereas task trainers only provide physical feedback for procedural skills. Furthermore, standardized patients offer realistic encounters with patients; however, actors can vary in their portrayal of patients, which affects the consistency of the training. Cine-VR can standardize patient interactions while also allowing for variations in patient interactions to balance consistency with the flexibility to adapt to human unpredictability. Altogether, cine-VR incorporates elements from manikins, advanced patient simulators, standardized patients, and task trainers into a single comprehensive training tool [33].

Our cine-VR training program has several advantages compared with other simulation approaches. First, our cine-VR was standardized, safe, and reusable. If needed, participants can repeat the training as often as desired to learn or reinforce knowledge and clinical skills. In addition, participants could practice difficult, high-stakes conversations through our "guided simulations," without any risk or harm to the patient, Mr Chen. Next, the multisensory experience of cine-VR engages emotional learning. Research on emotional learning shows that it increases long-term memory as well as increased acquisition of skills [34-36]. Finally, cine-VR training offers participants a glimpse into the lives of their patients, thereby increasing their awareness of another person's feelings and experiences. This form of perspective-taking increases empathy similar to the findings in our pilot study [32,37,38]. A 2017 study by Schutte and Silinovic [39] compared the impact of VR and traditional didactic learning on empathy among undergraduate students. They found participants in the VR group reported greater increases in empathy compared to the didactic presentation group. Thus, cine-VR has the potential to increase empathy in health professional students; more research is needed to assess empathy and other prosocial behaviors (ie, helping, sharing, and comforting) in future research. Considering patients value empathy from their health care professionals, often as much or even more than clinical expertise, simulation training that enhances empathy is crucial [40,41].

Limitations

Limitations of this study included the small sample, data collected from one site, heterogeneity of health care specialties, selection bias, subject bias, and lack of a control group. We recruited a sample of 65 participants for the training, and of those, 48 completed the pre- and postassessment. While the sample of 48 participants was small, it was sufficient to pilot the measures for our pilot study. Next, we collected data from one university in the Midwest, which limits the generalizability of our findings to all health professional students. Furthermore, participants represented 6 different health professional fields. Geriatric care is inherently interprofessional and typically requires collaboration from multiple health care professionals. However, the small sample size across 6 health professional programs limits the ability to draw conclusions about the cine-VR training program in these fields. In addition, students who volunteered to participate may have been more willing or motivated to participate in this cine-VR training on care for older adults. Thus, the findings may be susceptible to subject bias and social desirability bias. Finally, we did not include a control condition as a comparison group. Future research with a large sample size must include a control condition to determine the effectiveness of the cine-VR training. In future work, we plan to conduct a randomized controlled trial comparing the effectiveness of the cine-VR training to simulated patients. Simulated patients are used extensively in medical training to enhance the educational experience of health professional students; however, simulated patients can be costly and time-consuming. If we demonstrate that the cine-VR training is more effective or equally effective as simulated patients, cine-VR training may be a more sustainable, cost-effective approach to refine health professional students' clinical and communication skills without the risk of harming real patients.

Conclusions

We demonstrated that our cine-VR training program was feasible and acceptable to health professional students at our university in the Midwestern United States. Our findings suggest that the training increased awareness of discrimination towards people with disability, improved self-efficacy to identify and manage elder abuse and neglect, and increased empathy. More research is needed to confirm the effectiveness of this cine-VR training program on the complexity of care for older adults. Future research should use a randomized controlled trial to compare the cine-VR to an attention control condition, such as simulated patients. If confirmed, the cine-VR training program may be a new, effective approach to learning about multimorbidity, geriatric syndromes, and biases related to aging.

Acknowledgments

The research was funded by the Osteopathic Heritage Foundation Ralph S Licklider, DO Endowed Professor awarded to EAB. The creation of the virtual reality simulations was part of the Medicaid Care Experience Simulation funded by the Ohio Department of Medicaid and administered by the Ohio Colleges of Medicine Government Resource Center. The views expressed in the virtual reality simulations of this publication are solely those of the creators and do not represent the views of the state of Ohio or federal Medicaid programs.

Authors' Contributions

EAB, SM, ML, and CL provided substantial contributions to conception and design, acquisition of data, and analysis and interpretation of data. EAB, SM, ML, and CL drafted the article or revised it critically for important intellectual content. EAB, SM, ML, and CL gave final approval of the version of the article to be published. EAB, SM, ML, and CL agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

References

1. Mather M, Scommegna P. Fact sheet: aging in the United States. : Population Reference Bureau; 2024 URL: https://www.prb.org/resources/fact-sheet-aging-in-the-united-states/
2. Tavares JL, Cohen MA, Silberman S, Popham L. Chronic inequities: measuring disease burden among older adults in the US: a health and retirement study analysis.: National Council on Aging; 2022.
3. Stierman B, Afful J, Carroll MD, et al. National Health and Nutrition Examination Survey 2017-March 2020 Prepandemic Data Files-Development of Files and Prevalence Estimates for Selected Health Outcomes. Natl Health Stat Report 2021 Jun 14(158). [doi: 10.15620/cdc:106273] [Medline: 39380201]
4. Siew AL, Wong JW, Chan EY. Effectiveness of simulated patients in geriatric education: A scoping review. Nurse Educ Today 2021 May;100:104856. [doi: 10.1016/j.nedt.2021.104856] [Medline: 33740706]
5. Urtamo A, Jyväkorpi SK, Strandberg TE. Definitions of successful ageing: a brief review of a multidimensional concept. Acta Biomed 2019 May 23;90(2):359-363. [doi: 10.23750/abm.v90i2.8376] [Medline: 31125022]
6. Noto S. Perspectives on aging and quality of life. Healthcare (Basel) 2023 Jul 26;11(15):2131. [doi: 10.3390/healthcare11152131] [Medline: 37570372]
7. Sanford AM, Morley JE, Berg-Weger M, et al. High prevalence of geriatric syndromes in older adults. PLoS One 2020;15(6):e0233857. [doi: 10.1371/journal.pone.0233857] [Medline: 32502177]
8. Arai H, Chen LK. Aging populations and perspectives of geriatric medicine in Japan. Glob Health Med 2024 Feb 29;6(1):1-5. [doi: 10.35772/ghm.2024.01001] [Medline: 38450117]
9. O'Rourke LA, Morrison M, Grimsley A, Cotter VT. High-fidelity simulation and nurse clinical competence-an integrative review. J Clin Nurs 2023 May;32(9-10):1549-1555. [doi: 10.1111/jocn.16028] [Medline: 34453385]
10. Gaba DM. The future vision of simulation in health care. Qual Saf Health Care 2004 Oct;13 Suppl 1(Suppl 1):i2-10. [doi: 10.1136/qhc.13.suppl_1.i2] [Medline: 15465951]
11. Herrera-Aliaga E, Estrada LD. Trends and innovations of simulation for twenty first century medical education. Front Public Health 2022;10:619769. [doi: 10.3389/fpubh.2022.619769] [Medline: 35309206]
12. Sapkaroski D, Mundy M, Dimmock MR. Virtual reality versus conventional clinical role-play for radiographic positioning training: A students' perception study. Radiography (Lond) 2020 Feb;26(1):57-62. [doi: 10.1016/j.radi.2019.08.001] [Medline: 31902456]
13. Beverly E, Rigot B, Love C, Love M. Perspectives of 360-degree cinematic virtual reality: interview study among health care professionals. JMIR Med Educ 2022 Apr 29;8(2):e32657. [doi: 10.2196/32657] [Medline: 35486427]
14. Beverly EA, Love C, Love M, Williams E, Bowditch J. Using virtual reality to improve health care providers' cultural self-efficacy and diabetes attitudes: pilot questionnaire study. JMIR Diabetes 2021 Jan 27;6(1):e23708. [doi: 10.2196/23708] [Medline: 33502335]
15. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
16. Power MJ, Green AM, WHOQOL-DIS Group. The Attitudes to Disability Scale (ADS): development and psychometric properties. J Intellect Disabil Res 2010 Sep;54(9):860-874. [doi: 10.1111/j.1365-2788.2010.01317.x] [Medline: 20712699]
17. Simmons J, Wenemark M, Ludvigsson M. Development and validation of REAGERA-P, a new questionnaire to evaluate health care provider preparedness to identify and manage elder abuse. BMC Health Serv Res 2021 May 19;21(1):473. [doi: 10.1186/s12913-021-06469-2] [Medline: 34006271]
18. Hojat M, Gonnella J, Maxwell K. Jefferson Scales of Empathy (JSE) Professional Manual & User 's Guide: Jefferson Medical College; 2009.

19. Grassini S, Laumann K, Rasmussen Skogstad M. The use of virtual reality alone does not promote training performance (but sense of presence does). Front Psychol 2020;11:1743. [doi: 10.3389/fpsyg.2020.01743] [Medline: 32765384]
20. Slater M, Wilbur S. A Framework for Immersive Virtual Environments (FIVE): speculations on the role of presence in virtual environments. Presence: Teleoperators & Virtual Environments 1997 Dec;6(6):603-616. [doi: 10.1162/pres.1997.6.6.603]
21. LeDoux JE. Emotion circuits in the brain. Annu Rev Neurosci 2000;23:155-184. [doi: 10.1146/annurev.neuro.23.1.155] [Medline: 10845062]
22. Levine AI, Demaria S, Schwartz AD, Sim AJ. The Comprehensive Textbook of Healthcare Simulation: Springer; 2013.
23. Elendu C, Amaechi DC, Okatta AU, et al. The impact of simulation-based training in medical education: A review. Medicine (Baltimore) 2024 Jul 5;103(27):e38813. [doi: 10.1097/MD.0000000000038813] [Medline: 38968472]
24. Okuda Y, Bryson EO, DeMaria S Jr, et al. The utility of simulation in medical education: what is the evidence? Mt Sinai J Med 2009 Aug;76(4):330-343. [doi: 10.1002/msj.20127] [Medline: 19642147]
25. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
26. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. Acad Med 2003 Aug;78(8):783-788. [doi: 10.1097/00001888-200308000-00006] [Medline: 12915366]
27. Cooper JB, Taqueti VR. A brief history of the development of mannequin simulators for clinical education and training. Postgrad Med J 2008 Nov;84(997):563-570. [doi: 10.1136/qshc.2004.009886] [Medline: 19103813]
28. Gordon JA, Wilkerson WM, Shaffer DW, Armstrong EG. "Practicing" medicine without risk: students' and educators' responses to high-fidelity patient simulation. Acad Med 2001 May;76(5):469-472. [doi: 10.1097/00001888-200105000-00019] [Medline: 11346525]
29. Hillier M, Williams TL, Chidume T. Standardization of Standardized Patient Training in Medical Simulation: StatPearls Treasure Island (FL); 2025.
30. May W, Park JH, Lee JP. A ten-year review of the literature on the use of standardized patients in teaching and learning: 1996-2005. Med Teach 2009 Jun;31(6):487-492. [doi: 10.1080/01421590802530898] [Medline: 19811163]
31. Beverly EA, Love C, Love M, Lammert L, Bowditch J. Cinematic virtual reality for educating health care providers about type 2 diabetes, disability, and elder abuse and neglect: a pilot study. J Diabetes Sci Technol 2023 Sep;17(5):1160-1171. [doi: 10.1177/19322968231171586] [Medline: 37114917]
32. McCalla MM, Jones D, Grice R, et al. Feasibility of a cinematic-virtual reality training program for health professional students: a single-arm pre-post study. J Diabetes Sci Technol 2023 Sep;17(5):1181-1189. [doi: 10.1177/19322968231171136] [Medline: 37138539]
33. The Power of Virtual Reality Cinema for Healthcare Training: A Collaborative Guide for Medical Experts and Media Professionals: Routledge; 2022.
34. Tyng CM, Amin HU, Saad MNM, Malik AS. The influences of emotion on learning and memory. Front Psychol 2017;8:1454. [doi: 10.3389/fpsyg.2017.01454] [Medline: 28883804]
35. Steadman RH, Coates WC, Huang YM, et al. Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. Crit Care Med 2006 Jan;34(1):151-157. [doi: 10.1097/01.ccm.0000190619.42013.94] [Medline: 16374169]
36. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
37. Tong X, Gromala D, Kiaei Ziabari SP, Shaw CD. Designing a virtual reality game for promoting empathy toward patients with chronic pain: feasibility and usability study. JMIR Serious Games 2020 Aug 7;8(3):e17354. [doi: 10.2196/17354] [Medline: 32763883]
38. Herrera F, Bailenson J, Weisz E, Ogle E, Zaki J. Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. PLoS One 2018;13(10):e0204494. [doi: 10.1371/journal.pone.0204494] [Medline: 30332407]
39. Schutte NS, Stilinović EJ. Facilitating empathy through virtual reality. Motiv Emot 2017 Dec;41(6):708-712. [doi: 10.1007/s11031-017-9641-7]
40. Hirsch EM. The role of empathy in medicine: a medical student's perspective. Virtual Mentor 2007 Jun 1;9(6):423-427. [doi: 10.1001/virtualmentor.2007.9.6.medu1-0706] [Medline: 23218048]
41. Halpern J. What is clinical empathy? J Gen Intern Med 2003 Aug;18(8):670-674. [doi: 10.1046/j.1525-1497.2003.21017.x] [Medline: 12911651]

Abbreviations

cine-VR: cinematic–virtual reality
VR: virtual reality

Edited by H LaMonica; submitted 22.07.24; peer-reviewed by J Wardian, X Jing, Y Weng; revised version received 21.01.25; accepted 21.01.25; published 12.02.25.

Please cite as:
Beverly EA, Miller S, Love M, Love C
Feasibility of a Cinematic–Virtual Reality Program Educating Health Professional Students About the Complexity of Geriatric Care: Pilot Pre-Post Study
JMIR Aging 2025;8:e64633
URL: https://aging.jmir.org/2025/1/e64633
doi:10.2196/64633

Exploring Dance as a Therapeutic Approach for Parkinson Disease Through the Social Robotics for Active and Healthy Ageing (SI-Robotics): Results From a Technical Feasibility Study

Roberta Bevilacqua[1], MScPH; Elvira Maranesi[1], PhD; Marco Benadduci[1], BPT; Gabriella Cortellessa[2], PhD; Alessandro Umbrico[2], PhD; Francesca Fracasso[2], PhD; Giovanni Melone[3], MSc; Arianna Margaritini[1], MScPH; Angela La Forgia[3], MSIS; Pierpaolo Di Bitonto[4], BCS; Ada Potenza[5], MSc; Laura Fiorini[6,7], PhD; Carlo La Viola[6], PhD; Filippo Cavallo[6,7], PhD; Alessandro Leone[8], BSc; Andrea Caroppo[8], MSCE; Gabriele Rescio[8], PhD; Mauro Marzorati[9], PhD; Amedeo Cesta[2], PhD; Giuseppe Pelliccioni[10], MD; Giovanni Renato Riccardi[11], MD; Lorena Rossi[1], MEng

[1]Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

[10]Clinical Unit of Neurology, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA), Ancona, Italy

[11]Clinical Unit of Physical Rehabilitation, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA), Ancona, Italy

[2]National Research Council–Institute of Sciences and Technologies for Cognition, Rome, Italy

[3]Exprivia S.p.a., Innovation Lab, Innovation, Marketing & Technology, Molfetta, Italy

[4]Department of Soil, Plant and Food Science, University of Bari Aldo Moro, Bari, Italy

[5]Grifo multimedia Srl, Bari, Italy

[6]Department of Industrial Engineering, University of Florence, Florence, Italy

[7]The BioRobotics Institute, Scuola Superiore Sant'Anna, Pontedera, Italy

[8]National Research Council of Italy – Institute for Microelectronics and Microsystems, Lecce, Italy

[9]Institute of Biomedical Technologies, National Research Council, Segrate, Italy

Corresponding Author:
Elvira Maranesi, PhD
Scientific Direction, IRCCS INRCA, Via Santa Margherita 5, Ancona, Italy

Abstract

Background: Parkinson disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms. Recently, dance has started to be considered an effective intervention for people with PD. Several findings in the literature emphasize the necessity for deeper exploration into the synergistic impacts of dance therapy and exergaming for PD management. Moreover, socially engaging robotic platforms equipped with advanced interaction and perception features offer potential for monitoring patients' posture and enhancing workout routines with tailored cues.

Objective: This paper presents the results of the Social Robotics for Active and Healthy Ageing (SI-Robotics) project, aimed at designing an innovative rehabilitation program targeted at seniors affected by (early-stage) PD. This study therefore aims to assess the usefulness of a dance-based rehabilitation program enriched by artificial intelligence–based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. The acceptability of the system is also investigated.

Methods: The study is designed as a technical feasibility pilot to test the SI-Robotics system. For this study, 20 patients with PD were recruited. A total of 16 Irish dance–based rehabilitation sessions of 50 minutes were conducted (2 sessions per week, for 8 wks), involving 2 patients at a time. The designed rehabilitation session involves three main actors: (1) a therapist, (2) a patient, and (3) a socially interacting robot. To stimulate engagement, sessions were organized in the shape of exergames where an avatar shows patients the movements they should perform to correctly carry out a dance-based rehabilitation exercise.

Results: Statistical analysis reveals a significant difference on the Performance-Oriented Mobility Assessment scale, both on balance and gait aspects, together with improvements in Short Physical Performance Battery, Unified Parkinson Disease Rating Scale–III, and Timed Up and Go test, underlying the usefulness of the rehabilitation intervention on the motor symptoms of PD. The analysis of the Unified Theory of Acceptance and Use of Technology subscales provided valuable insights into users' perceptions and interactions with the system.

Conclusions: This research underscores the promise of merging dance therapy with interactive exergaming on a robotic platform as an innovative strategy to enhance motor function, balance, gait, and overall quality of life for patients grappling with PD.

Trial Registration: ClinicalTrials.gov NCT05005208; https://clinicaltrials.gov/study/NCT05005208

(JMIR Aging 2025;8:e62930) doi:10.2196/62930

KEYWORDS

Parkinson disease; rehabilitation; Irish dancing; balance; gait; socially interacting robot

Introduction

Parkinson disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms, such as tremors, bradykinesia, rigidity, and postural instability, as well as nonmotor symptoms, including cognitive impairment, mood disorders, and autonomic dysfunction [1]. These symptoms can significantly impact patients' quality of life and independence as the disease progresses. While pharmacological interventions, such as dopamine replacement therapy, remain the mainstay of PD management, their efficacy may diminish over time and they often fail to address the nonmotor symptoms adequately [2]. Physical therapy represents the key adjuvant treatment for PD, offering potential benefits in improving motor function, balance, gait, and overall mobility [3]. Among various forms of physical therapy, dance-based interventions have gained attention due to their multifaceted approach targeting motor, cognitive, and psychosocial aspects of the disease [4]. Dance therapy engages patients in rhythmic movements, music, and social interaction, promoting coordination, flexibility, and emotional well-being [5].

Exergaming, which involves interactive video games requiring physical movement, has emerged as a novel approach to delivering dance-based therapy in both home-based and clinical settings [6]. Exergames offer the advantages of personalized, adaptable, and enjoyable exercise programs, potentially enhancing patient motivation and adherence [7]. Recent research has highlighted the potential of dance therapy and exergaming to address the complex needs of PD patients. A randomized controlled trial by Duncan and Earheart [4] demonstrated significant improvements in motor function and balance among patients with PD participating in community-based dancing compared with a control group. Furthermore, a study by Barry et al [7] found that exergaming interventions were effective in improving mobility and reducing fall risk in patients with PD.

These findings underscore the importance of further investigating the combined effects of dance therapy and exergaming in PD management. Furthermore, the reliability and performance of current sensing technologies provide physiological data that are useful to monitor the state and the quality of the exercises objectively performed by patients. For example, commercial wearable devices could be easily integrated during rehabilitation sessions to gather information about the heart or breath rate of involved patients and objectively evaluate the metabolic effort. Socially interactive robotic platforms with their interaction and perception capabilities (eg, onboard cameras) are well suited to monitor the posture of patients and support the execution of the exercises through tailored stimuli. A more comprehensive approach may integrate advanced technologies, such as sensor technology for real-time monitoring, artificial intelligence (AI) techniques for personalized difficulty adjustments, and robotic platforms to deliver personalized feedback and enhance the rehabilitation experience.

In this context, the Social Robotics for Active and Healthy Ageing (SI-Robotics) project designed an innovative rehabilitation program targeted at seniors affected by (early stage) PD [8]. The approach relies on the integration of several AI technologies ranging from knowledge representation and reasoning, for user modeling and personalization to machine learning and automated planning (AP) for continuous proactive and adaptive assistance [9]. Leveraging the mentioned interventions based on different types of dance [10,11], the idea is to develop a technology-enhanced, dance-based rehabilitation program where a socially interacting robot and several sensing devices (wearable sensors and 3D cameras mainly) support therapists and patients during the execution of the exercises. This study therefore aims to assess the usefulness of a dance-based physical therapy program enriched by AI-based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. We hypothesized that the proposed technology-enhanced dance therapy would lead to greater improvements in PD symptoms.

Methods

The study is designed as a technical feasibility pilot to test the SI-Robotics system. The entire protocol, including the description of scales, the platform functioning, the training program, and procedures has been previously described in detail [8].

Subjects

A total of 20 patients with PD were selected by the outpatient department at the Clinical Unit of Physical Rehabilitation, Istituto di Ricovero e Cura a Carattere Scientifico Istituto Nazionale Ricovero e Cura per Anziani (IRCCS INRCA). The patients were recruited if they were over 65 years old; able to provide informed consent; had a stage of Hoen and Yahr scale between 1 and 2 [12]; had a Functional Ambulation Category score ≥2 [13]; had a Rankin Scale score ≤3 [14]; had stability of drug treatment for at least 1 month; had a Geriatric Depression Scale 4-item score ≤1 [15]; had a Mini-Mental State Examination ≥24 [16]; and could maintain an upright posture≥30 seconds, evaluated by a trained physiotherapist during the recruitment. The evaluation of compliance with the inclusion and exclusion criteria was performed during the recruitment session. Once we completed this phase, informed consent was obtained. The patient's clinical assessment was performed at the start and the end of the treatment. In particular, the evaluation consisted of the administration of the following scale: measurement of functional state with the Barthel Index [17]; physical performance with Tinetti's Performance-Oriented

Mobility Assessment (POMA) [18], the Short Physical Performance Battery (SPPB) [19], the 6-Minute Walking Test [20], and the Timed Up and Go test (TUG) [21]; evaluation of the quality of life with the 12-Item Short-Form Health Survey (SF-12) [22]; fear of falling with Falls Efficacy Scale-International (FES-I) [23]; and the assessment of the prognosis of PD with the Unified Parkinson Disease Rating Scale – III (UPDRS-III) [24].

Intervention

Figure 1 shows the experimental setting in a protected environment (gymnasium) and the positioning of the technologies, the patient, and the physiotherapist, during the dance-based rehabilitation sessions.

Figure 1. Technological components and actors of Social Robotics for Active and Healthy Ageing (SI-Robotics) rehabilitation sessions.

A total of 16 therapy sessions of 50 minutes were conducted (2 sessions per wk, for 8 wks), involving 2 patients at a time. Cardiac and respiratory activity was monitored during dancing treatments to detect heart rate and breathing frequency. The participants were required to complete at least 80% of the sessions. Each session involved the following activities: (1) breathing, relaxation, and postural harmonization exercises (5 min); (2) active mobility and stretching exercises (5 min); (3) the SI-Robotics intervention (35 min) consisting of the AI-enhanced “Let’s Dance” game, and a socially interacting robot monitoring the execution of the physical exercises of patients; and (4) relaxation exercises (5 min).

After selecting a difficulty level, players are presented as dancers on the screen, along with footprints that suggest the movement to be done. Each task can vary from simple aerobic exercises (side steps, arm raising, hand clapping, etc) to choreographies. At the end of the game session, a score is presented to the patient. This would make a patient aware of the quality of the performance. This score can also be used by the therapist to assess the patient’s performance and the possible increase in the difficulty level of the game.

The next sections describe the technological components of the session in more detail.

The AI-Based Rehabilitation Session

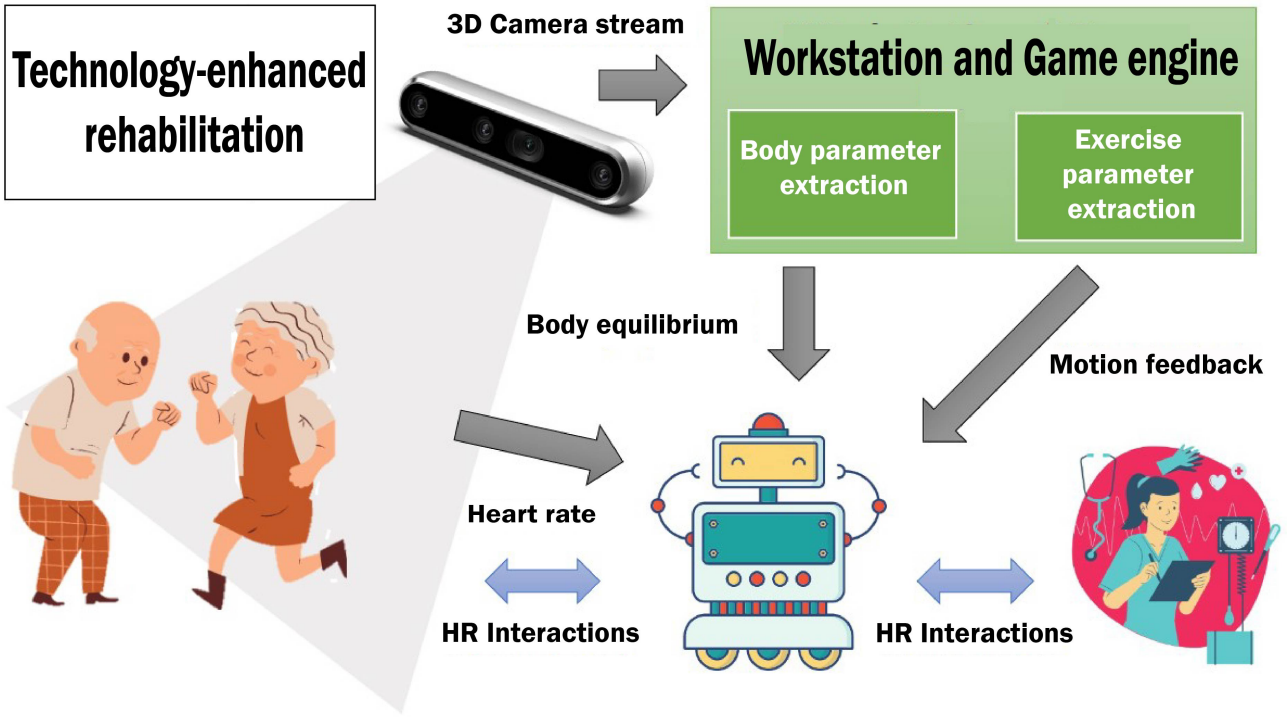
The designed rehabilitation session involves three main actors: (1) a therapist, (2) a patient, and (3) a socially interacting robot. To stimulate engagement, sessions are organized in the shape of exergames where an avatar shows patients the movements they should perform to correctly carry out a rehabilitation exercise (eg, dancing steps). Given a set of data about the needs of a group of patients and about stimulation capabilities of known exercise, a game engine integrates an AI planner [25] and synthesizes exercises by selecting a (sub)set of movements (ie, dancing steps) that best fit the clinical objectives of the session (ie, of the current patient). The exergame is enriched with AP to synthesize suitable physical exercises, contextualized to the clinical objectives of the rehabilitation sessions [25]. Many works in the literature investigate the use of AI in health care and PD [11,26] AI is used for example to predict the wearing-off of symptoms [27], support decisions [28], or early diagnosis of PD [29]. The majority of these works adopt AI solutions based on deep or machine learning and focus on the diagnosis of the disease. Few works investigate the use of AP to support therapists in the synthesis of rehabilitation programs. The study by Gonzalez et al [30], for example, integrates AP into a control architecture to allow a social robot to physically show motions to a patient during physical rehabilitation. Baschieri et al [31] uses AP to synthesize simulation scenarios within a serious game for cognitive rehabilitation. This work pursues an objective similar to ours but in a different clinical scenario.

As shown in Umbrico et al [25] SI-Robotics proposes AP to support the synthesis of physical rehabilitation programs for patients with PD . A key aspect of the developed planning framework is the combined reasoning about spatial and clinical effects of stimuli (ie, motions or dancing steps) needed to synthesize plans that are technically valid and effective from a clinical point of view. Compared with a manual definition of the rehabilitation session, the use of a planner aims to improve the quality, accuracy, and engagement of the resulting rehabilitation programs. A therapist provides a planner with data about the rehabilitation session (song time, song rhythm, and difficulty level) and the clinical objectives. Given this input, the planner decides a sequence of steps, optimized according to an objective function encoding the specified clinical objective. Planned steps are thus chosen according to the rehabilitation needs of the participating patient (personalization).

In addition, several devices enrich the session to extract useful data about the health conditions and the performance of patients.

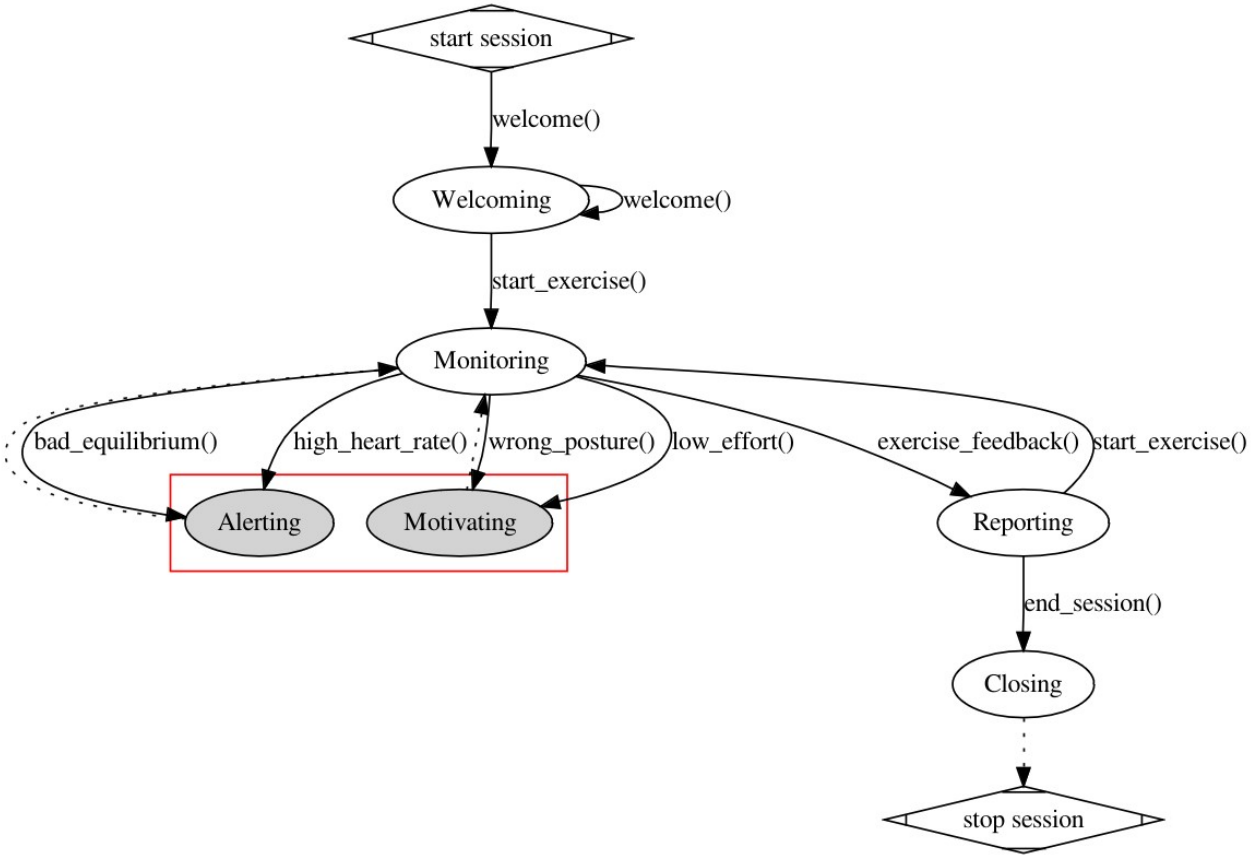
The generated data streams are gathered by the robot that embodies the AI-based reasoning modules. Overall, three types of perception components were considered: (1) wearable sensing devices (eg, sensorized shirts) that constantly produce data streams about physiological parameters of patients (eg, heart rate) [32]; (2) a video-processing component elaborating 3D camera data to extract kinematics about motions and produce data about body posture and body equilibrium of patients and; (3) a video-processing component elaborating 3D camera data to analyze patients' behaviors and produce data about the correctness of performed motions (ie, feedback). Figure 2 below shows a conceptual overview of the designed session pointing out a subset of the main hardware and software components and the data streams generated during the execution of a rehabilitation session.

Figure 2. Automata describing the assistive behavior of the robot during a rehabilitation session. HR: Human-Robot



In addition, a socially interacting robot should support the execution of a session by recognizing the current state of a session according to data gathered from the environment, autonomously detecting events that would trigger different phases, and setting the behavioral goals needed to support a session. The robotic platform consists of a novel social robotic platform designed within SI-Robotics and developed by Co-Robotics [33]. Figure 3 shows the lifecycle determining the assistive behavior of the robot during each rehabilitation session.

Figure 3. Architectural structure of the artificial intelligence–based cognitive control of the robot.



A rehabilitation session starts with a welcoming state where the robot introduces patients to the session. Patients do not know the structure of the rehabilitation and are not familiar with the involved technology. The objective of this state is therefore to prepare patients for the therapy by explaining the structure of the session and the technology used. When all patients are ready to start a session, the robot enters into a "monitoring" state gathering data from the environment to monitor the performance and health state of patients. When a rehabilitation exercise ends the robot enters into the "reporting" state to interact with the therapist to comment on the exercise and the observed performance. This state is especially useful to enrich the subjective experience of the therapist by enhancing the awareness of the therapist with objective feedback on the quality of the exercises. It is indeed meant as a support for the decisions on how to proceed with the session (eg, change or repeat the exercises, difficulty of the next exercise, etc). If the therapist decides to start a new exercise, then the robot enters again into the monitoring state. Otherwise, the robot enters into the "closing" state to finalize the session. In this latter case, the robot interacts with patients to ask their personal feelings about the session and to show a qualitative assessment of their session.

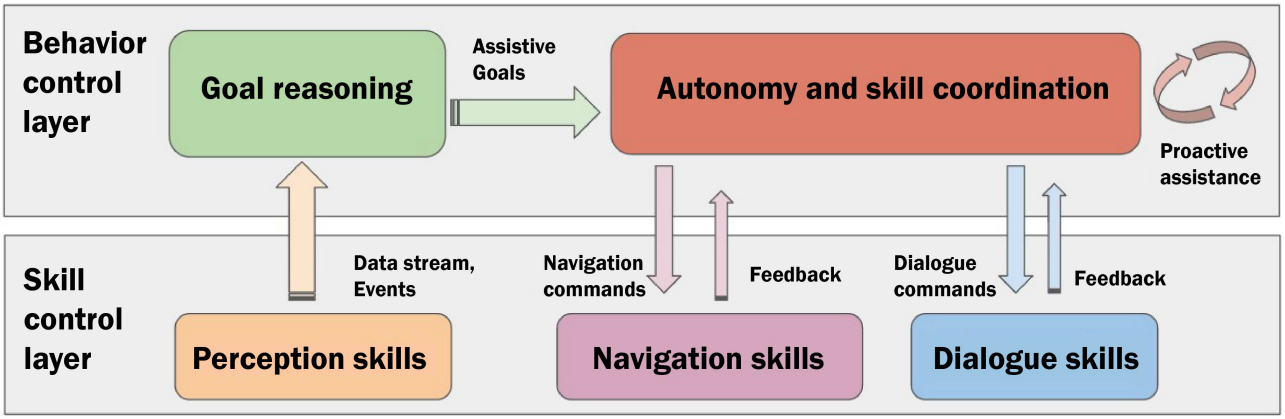
Different types of interaction are foreseen during an exercise, depending on the type of situation recognized by the robot. Two types of interactions have been considered characterized by the motivating and the alerting states in the automata. The motivating phase allows the robot to play the role of a coach during the session stimulating patients to perform better or correcting their behavior. The robot is supposed to enter the motivating phase when it detects a heart rate below some minimum "training threshold" or when it detects a wrong posture of the body, for example, bad alignment or asymmetrical motion of lower or upper limbs.

The alerting phase allows the therapist (and the robot) to intervene when some critical situation concerning the health state of a patient is detected during the session. In such case, the robot is supposed to warn the therapist about the danger and simultaneously interact with the involved patient, even interrupting the exercise if necessary. Specifically, the robot enters the alerting phase when it detects the risk of a fall or an anomalous heart rate for a patient. Both these situations are critical for the considered target of patients and need a prompt intervention.

AI-Based Control of Robotic Skills

A key aspect (and still an open challenge) in the design of autonomous robotic agents is the integration of different AI technologies and Robotics [34,35]. SI-Robotics proposes an AI-based control architecture with twofold objectives: (1) to support the abstraction and reasoning capabilities needed to recognize health-related situations and assistive objectives and (2) to coordinate robotic skills to "act" in the environment and (autonomously) achieve contextualized assistive objectives. Figure 4 shows an overview of the designed architecture with the main functional components involved in the control process. Two main architectural layers are considered and are organized according to a cognitive architecture inspired by the Dual Process Psychological Theory [36].

Figure 4. Designed architecture.



A skill control layer encapsulates the control components responsible for providing the robot with the "low-level" capabilities needed to perceive the environment and interact with the environment. The Navigation Skills component encapsulates functionalities that allow a robot to robustly navigate within the environment and autonomously reach known locations. The Dialog Skills encapsulate natural language processing functionalities that allow a robot to talk with patients and therapists. This component in particular enables an affective interaction [37] by adapting the way the robot talks with a patient to the detected emotional state and known personality traits.

While some level of perception is needed for both navigation and dialogue skills, the Perception Skills component emphasizes the additional capability needed to process data from deployed environmental devices. It encapsulates low-level data processing to provide the higher decisional level with refined information. A behavior control layer supports AI-based control features to contextualize observations, trigger assistive goals, and coordinate underlying skills to implement the desired assistance. The Autonomy and Skill Coordination component encapsulates an AI plan–based controller that supports behavior continuity coordinating robotic skills. The Goal Reasoning component instead is in charge of contextualizing observations to automatically set assistive goals. The combination of goal reasoning and a plan-based controller is the key architectural feature enabling proactive and autonomous goal-oriented behaviors of the robot [38].

The AI-based architectural elements composing the coordination layer of Figure 4 have been embedded into the robotic platform and integrated through the Robot Operating System (ROS) using the ROS Melodic distribution [39]. The coordination component responsible for the synthesis of the assistive behavior has been implemented in the shape of an AI plan–based controller relying on AP. The component uses the open-source framework ROXANNE [40], which implements goal-oriented acting capabilities into an ROS [41]. ROXANNE allows the robot to receive high-level assistive goals and (autonomously) synthesize and execute navigation and dialogue commands or tasks to realize a desired assistive behavior. The goal reasoning component correlates data received from the perception component with the autonomous behavior of the robot. It endows the robot with the cognitive capabilities needed to understand the evolution of a rehabilitation session autonomously select suitable planning goals and consequently decide the assistance needed in a particular context.

Embedded AP models the behavioral constraints needed to correctly support a rehabilitation session. Behavioral constraints are specified in the shape of temporal constraints and require interacting tasks performed through available skills. In this regard, Table 1 shows the modeled assistive goals and the parameters defined to contextualize interactions. Furthermore, the table briefly describes the behavioral constraints specified in the model to correctly implement the desired assistance and support the associated goal. It is worth noting that AI planning facilitates robot programming since it supports an "easy specification" of desired behavioral dynamics without hard-coding the implemented skills.

Table . Acting goals supported by the robot-embedded controller to synthesize assistive behaviors during a rehabilitation session.

Goal	Interaction parameters	Interaction parameters
Welcoming	• Therapist: string • Patient: string • Gender: string • Novel: boolean	The robot moves from its current location to the welcoming area to call the specified patient and take patient to the dressing area to wear the sensorized shirt. Then it guides the patient to the rehabilitation area to start the session. If the patient is new to the session (flag novel set to true) the robot explains the organization of the rehabilitation and the functioning of the devices. Otherwise, the robot shows data about the last session to motivate the patient for the current one.
Exercise start	__[a]	When an exercise begins the robot starts moving around the rehabilitation area and monitoring patients' state by collecting physiological, body posture, and performance data.
Exercise report	• Patient: string • Gender: string • Score: double • Past_score: double • Global_score: double	When an exercise ends the robot moves close to a patient to show a brief report about the patient performance. The robot shows the achieved score and (if available) compares it to the average score of the patient in the past and the average score of other patients. This would make a patient aware of the quality of the performance.
Low heart rate and high heart rate	• Therapist: string • Patient: string • Gender: string • Heart_rate: double	The robot periodically analyzes trends of gathered data about the heart rate of a patient. Thresholds are parametric and computed according to the age of the patient as follows: HR_max=80% (220 – age) HR_min=40% h_max The objective of the session is to keep the heart rate of patients within HR_max and HR_min to stimulate a proper metabolic effort. In the case of low heart rate (i.e., value below HR_min) the robot implements a motivational behavior asking a patient to improve the performance. In case of high heart rate (i.e., value above HR_max), the root implements an assistive behavior alerting the therapist and asking the patient to reduce the performance
Bad posture	• Patient: string • Gender: string • Tilt: integer in [0, 1] • Arms: integer in [–1, 1] • Legs: integer in [–1, 1]	The robot constantly monitors the motions and the inclination of patients. If anomalous kinematics parameters are detected, for example, wrong inclinations of the body or wrong alignment of arms or legs, the robot implements a motivational behavior to correct the body posture of a patient and thus stimulate correct motions.
Bad equilibrium	• Therapist: string • Patient: string • Gender: string • Lateral: integer in [–1,1] • Frontal: integer in [–1,1]	In addition to the body posture, the robot constantly monitors the equilibrium of patients to prevent dangerous falls during physical therapy. If dangerous equilibrium conditions are detected from camera data (eg, the center of mass is outside the patient's base) the robot implements an assistive behavior to promptly notify the therapist and warn the patient about the danger or falling.

[a]Not applicable.

Outcomes

All assessment procedures adhere to a standardized protocol. Specifically, the primary focus of the study revolves around enhancing balance, and gait, and alleviating the fear of falling among elderly patients with PD. This is gauged through the utilization of the 3 POMA scales (POMA balance, POMA Gait, and POMA Total) after the 10 treatment sessions, as a result of the Irish dance intervention. In addition, secondary measures include evaluating the gait speed of elderly patients with PD, their fear of falling (assessed via FES-I), their physical

performance (SPPB), their autonomy in daily living activities (UPDRS-III), and their overall physical and psychological well-being (SF-12) together with the evaluation of the Unified Theory of Acceptance and Use of Technology (UTAUT) that is composed by 10 subscales (anxiety; attitude; facilitating conditions; intention to use; perceived adaptability; perceived enjoyment; perceived ease of use; perceived usefulness; social influence; and trust).

Statistical Analysis

Continuous variables were presented as either mean and SD or median and IQR, depending on their distribution, which was determined using the Kolmogorov-Smirnov test. Categorical variables were expressed as absolute numbers and percentages. To test statistically significant differences ($P<.050$) between pre- and postconditions, Wilcoxon signed-rank test (for non-normal distribution) or paired test (for normal distribution) were used, alongside simple descriptive statistics such as means, medians, and SDs as appropriate. The statistical analysis was performed using SPSS software.

Ethical Considerations

The study was approved by the Ethics Committee of the Istituto Nazionale Ricovero e Cura per Anziani, (IRCCS INRCA) on June 17, 2021 (CE INRCA 21004). The protocol is registered on ClinicalTrials.gov with trial registration number NCT05005208 (October 23 , 2023). All participants signed the informed consent and data processing consent. The data are anonymised so that the identity of the subject cannot be traced. No compensation is provided for participation in the study.

Results

Demographic and clinical data of the patients are reported in Table 2. Two participants dropped out because they did not complete the rehabilitation program.

The CONSORT (Consolidated Standards of Reporting Trials) flowchart is shown in Figure 5.

Gender differences were not statistically significantly different in all scales used to select the sample.

Table . Baseline demographic and clinical profile.

		Total (n=18)	Male (n=10)	Female (n=8)	*P* value
Age (years), mean (SD)		75.3 (5.5)	73.7 (5.4)	77.2 (5.2)	.18
Marital status, n (%)					.48
	Married	15 (83.3)	9 (90)	6 (75)	
	Single	1 (5.5)	0 (0)	1 (12.5)	
	Widowed	2 (11.2)	1 (10)	1 (12.5)	
Educational level, n (%)					.17
	Primary education	4 (22.2)	1 (10)	3 (37.5)	
	Secondary education	10 (55.6)	6 (60)	4 (50)	
	University or more	4 (22.2)	3 (30)	1 (12.5)	
Hoehn and Yahr score, mean (SD)		1.8 (0.3)	1.9 (0.3)	1.7 (0.4)	.42
Rankin Scale score, mean (SD)		1.6 (0.9)	1.4 (0.8)	1.8 (1.1)	.33
GDS[a], mean (SD)		3.0 (0.7)	2.9 (0.7)	3.2 (0.7)	.33
FAC[b], mean (SD)		4.6 (0.6)	4.7 (0.4)	4.5 (0.74)	.50
MMSE[c], mean (SD)		29.4 (0.9)	29.7 (0.6)	29.1 (0.7)	.22
Barthel Index score, mean (SD)		95.5 (5.6)	94.5 (6.8)	96.8 (3.7)	.39

[a]GDS: Geriatric Depression Scale.

[b]FAC: Functional Ambulation Category.

[c]MMSE: Mini-Mental State Examination.

Figure 5. The CONSORT (Consolidated Standards of Reporting Trials) flowchart.

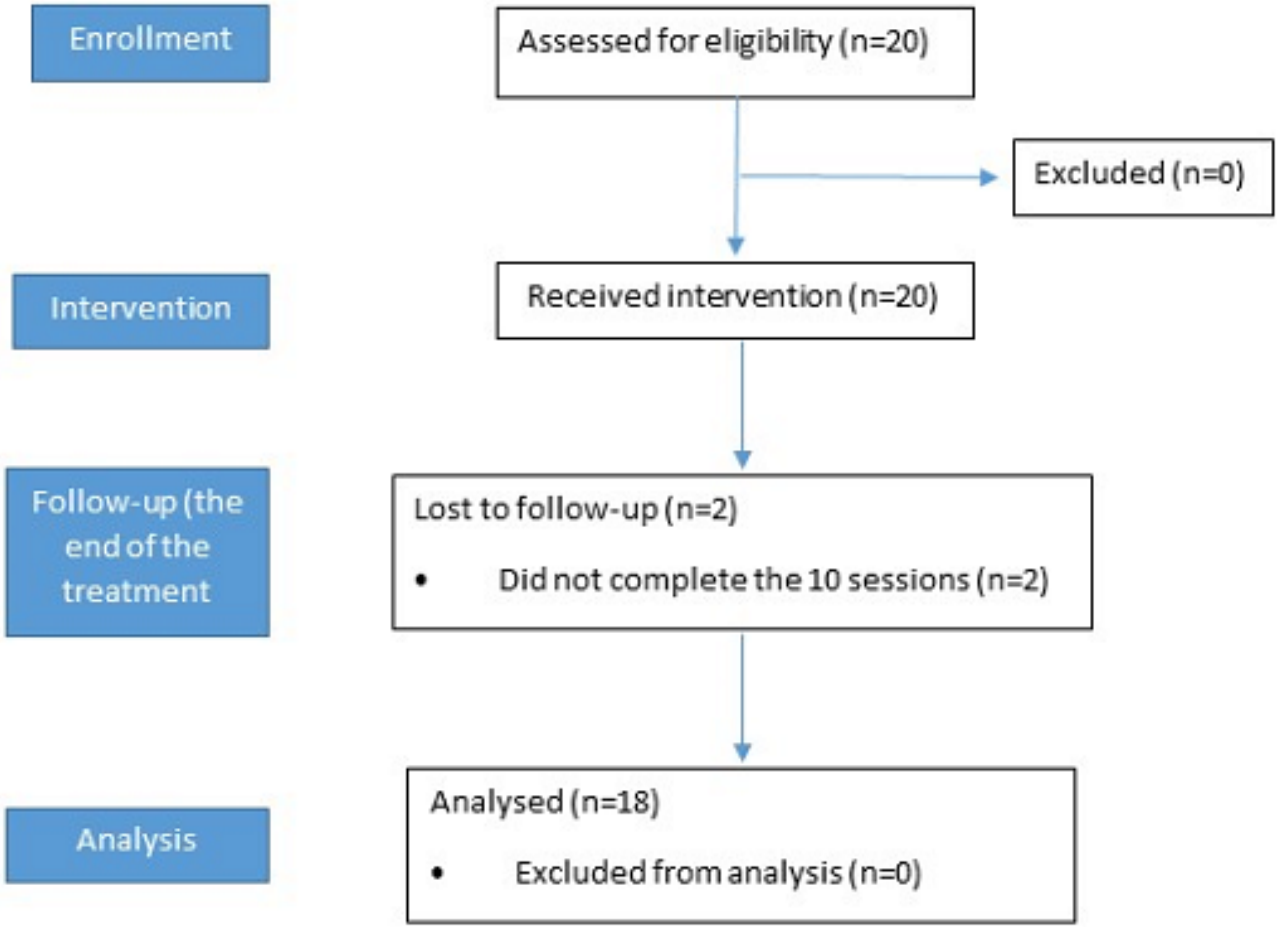


Table 3 shows differences between pre- and postintervention on the functional state scales with the UPDRS-III and the SPPB, gait and balance performance on Tinetti POMA Gait and POMA Balance, evaluation of quality of life with SF-12 and its subscores (physical component score and mental component score), fear of falling (FES-I) together with the TUG execution time, the meters covered during the 6-Minute Walking Test and the gait speed.

Statistical analysis revealed a significant effect on UPDRS-III and TUG together with improvements in the POMA scale, both on balance and gait aspects, and in SPPB. However, for the POMA and SPPB scales, there was a low Cohen *d* coefficient, underlining a probable effect of the low sample size on these results. In contrast, the effect size for UPDRS-III and TUG was high (>0.70), highlighting the usefulness of the rehabilitation intervention on the motor symptoms of PD also in our small sample.

Table . Mean (SD) values of the mean of pre- and postintervention scores on the UPDRS-III[a], SPPB[b], POMA[c] (total, gait and balance), SF-12[d] (total, physical, and mental component score), FES-I[e], TUG[f] execution time, 6MWT[g], and gait speed. Pre-post and between-group comparisons are reported for each score (P<.05). The effect size (Cohen d) is also reported.

		Preintervention scores, mean (SD)	Postintervention scores, mean (SD)	P value	Cohen d
UPDRS-III		13.56 (1.85)	13.83 (1.85)	.01[h]	−0.73
SPPB		8.83 (2.61)	10.39 (2.45)	.02[h]	0.10
POMA					
	POMA Total	24.11 (3.92)	26.28 (2.74)	.001[h]	−0.07
	POMA Gait	9.83 (2.30)	10.94 (1.55)	.01[h]	−0.02
	POMA Balance	14.28 (2.16)	15.33 (1.37)	.01[h]	−0.10
SF-12					
	SF-12-Tot[i]	31.28 (1.99)	32.33 (2.42)	.10	−0.20
	PCS-12[j]	13.56 (1.84)	13.83 (1.85)	.49	−0.23
	MCS-12[k]	17.72 (1.22)	18.50 (1.91)	.12	−0.16
FES-I		10.78 (1.92)	10.67 (2.22)	.83	−0.29
TUG (s)		11.0 (3.91)	9.19 (3.14)	.002[h]	−0.82
6MWT (m)		361.50 (99.33)	373.72 (98.67)	.40	−0.19

[a]UPDRS-III: Unified Parkinson Disease Rating Scale – III.

[b]SPPB: Short Physical Performance Battery.

[c]POMA: Performance Oriented Mobility Assessment.

[d]SF-12: 12-Item Short-Form Health Survey.

[e]FES-I: Falls Efficacy Scale–International.

[f]TUG: Timed Up and Go test.

[g]6MWT: 6-Minute Walking test.

[h]P values from matched-pairs Student t test.

[i]SF-12-Tot: 12-Item Short-Form Health Survey total score.

[j]PCS-12: 12-Item Short-Form Health Survey physical component score

[k]MCS-12: 12-Item Short-Form Health Survey mental component score.

Table 4 reports the scores of the subscales of UTAUT.

The analysis of the UTAUT subscales provided valuable insights into users' perceptions and interactions with the system. In particular, attitude, perceived adaptability, enjoyment, ease of use, usefulness, and social influence were generally positive.

Table . Mean (SD) of the mean of the Unified Theory of Acceptance and Use of Technology (UTAUT) subscales scores.

UTAUT subscales		Score, mean (SD)	Range
Anxiety (ANX)	Evoking anxious or emotional reactions when using the system.	7.4 (4.2)	0 - 20
Attitude (ATT)	Positive or negative feelings about the appliance of the technology	11.2 (2.5)	0 - 15
Facilitating conditions (FC)	Objective factors in the environment that facilitate using the system	2.6 (1.7)	0 - 15
Intention to use (ITU)	The outspoken intention to use the system over a longer period in time	3 (0)	0 - 10
Perceived adaptability (PAD)	The perceived ability of the system to be adaptive to the changing needs of the user.	10.5 (3.0)	0 - 15
Perceived enjoyment (PENJ)	Feelings of joy or pleasure associated by the user with the use of the system.	18.2 (2.2)	0 - 25
Perceived ease of use (PEOU)	The degree to which the user believes that using the system would be free of effort	18.1 (3.8)	0 - 25
Perceived usefulness (PU)	The perceived ability of the system to perform sociable behavior.	11.4 (2.9)	0 - 15
Social influence (SI)	The user's perception of how people who are important to him think about him using the system	8 (2.4)	0 - 10
Trust	The belief that the system performs with personal integrity and reliability.	7.5 (1.8)	0 - 10

Discussion

Principal Findings

This study aims to investigate the usefulness of the SI-Robotics intervention based on AI robotic assistance and exergame stimulation, in the context of PD. This study aims to assess the ability of a dance-based physical therapy program enriched by AI-based exergames and contextual robotic assistance in improving motor function, balance, gait, and quality of life in patients with PD. The findings of this study demonstrate the potential usefulness of a dance-based physical therapy program using a robotic platform to deliver exergames in improving motor function, balance, gait, and quality of life in patients with PD. Results suggest that integrating dance therapy with engaging technology may lead to greater improvements in PD symptoms compared with conventional physical therapy alone. Although the statistical analysis revealed significant improvements in motor function, balance, and gait (eg, UPDRS-III, POMA, and TUG scores), the effect sizes for several outcome measures were relatively small (Cohen d<0.2 for POMA and SPPB scales). This suggests that, while the observed changes were statistically significant, the practical impact of the intervention on certain motor and functional outcomes may be limited. These small effect sizes could be partly attributed to the small sample size, which may have reduced the power to detect larger effects.

This study adds to the growing body of research supporting the use of dance therapy and exergaming as adjunctive treatments for PD. Previous studies have shown that both dance therapy and exergaming can independently improve motor function, balance, and mobility in patients with PD [4,7]. By combining these approaches, our intervention aimed to target the multifaceted nature of PD symptoms, including motor, cognitive, and psychosocial aspects.

The results are consistent with previous findings demonstrating the benefits of dance therapy and exergaming in PD management. For example, Duncan and Earheart [4] reported significant improvements in motor function and balance among patients with PD participating in community-based dancing, while Barry et al [7] found that exergaming interventions were effective in improving mobility and reducing fall risk in patients with PD. These studies, along with ours, highlight the potential of integrating dance therapy with exergaming to address the complex needs of patients with PD.

One novel aspect of our intervention is the use of the SI-Robotics system, a robotics-based platform designed to engage patients with PD in Irish set dancing. This innovative approach combines the physical and cognitive benefits of dance therapy with the interactive and adaptable nature of exergaming. By incorporating personalized avatars and choreography, the "Let's Dance" game provided a stimulating and enjoyable exercise experience for participants, potentially enhancing motivation and adherence to the intervention.

Regarding the analysis of technology acceptance among the population studied, the results indicate a generally positive perception and interaction with technology. In particular, users displayed a positive attitude towards the technology, indicating

favorable feelings toward it. They also found the system adaptable, enjoyable, easy to use and useful showing that users experienced pleasure while using the system. Finally, trust in the system was moderately high, indicating that users believed it to be reliable and trustworthy. These highlight how the system successfully meets user expectations by offering an engaging, straightforward, and beneficial experience.

Our study focused on elderly patients with PD with mild to moderate disease severity, as reflected by their scores on the Hoen and Yahr scale, Functional Ambulation Category, Rankin Scale, Geriatric Depression Scale, and Mini-Mental State Examination. These inclusion criteria aimed to ensure that participants were physically and cognitively capable of engaging in the intervention safely and effectively. Future research could explore the applicability of our intervention to a broader range of patients with PD , including those with more advanced disease stages.

Limitations of our study include the small sample size and lack of a control group for comparison. While our results provide preliminary evidence of the usefulness of the intervention, larger-scale randomized controlled trials are needed to confirm these findings and establish the optimal dosage and timing of dance-based physical therapy with exergames for PD management. In addition, longer-term follow-up assessments could help evaluate the sustainability of the intervention effects over time.

Conclusions

In conclusion, this study highlights the potential of integrating dance therapy with exergaming integrated with an interactive robotic platform as a novel approach to improving motor function, balance, gait, and quality of life in patients with PD. The SI-Robotics intervention offers a promising avenue for future research and clinical practice in the management of PD symptoms. To strengthen the validity of the current findings and ensure their broader applicability, future research will focus on replicating the study with a larger sample. This will help to verify the generalizability of the results providing a more comprehensive understanding of the observed effects. Moreover, further studies are warranted to explore the long-term effects and feasibility of implementing this intervention in diverse clinical settings.

Acknowledgments

The study has received funding under the project Social Robotics for Active and Healthy Ageing (SI-Robotics (ARS01_01120), funded by the Italian Ministry of Education, Universities and Research, under the National Operational Program Area, Technologies for Living Environments.

Conflicts of Interest

None declared.

References

1. Postuma RB, Berg D, Stern M, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord 2015 Oct;30(12):1591-1601. [doi: 10.1002/mds.26424] [Medline: 26474316]
2. Olanow CW, Stern MB, Sethi K. The scientific and clinical basis for the treatment of Parkinson disease (2009). Neurology (ECronicon) 2009 May 26;72(21 Suppl 4):S1-136. [doi: 10.1212/WNL.0b013e3181a1d44c] [Medline: 19470958]
3. Tomlinson CL, Patel S, Meek C, et al. Physiotherapy versus placebo or no intervention in Parkinson's disease. Cochrane Database Syst Rev 2014;2014(6):CD002815. [doi: 10.1002/14651858.CD002815]
4. Duncan RP, Earhart GM. Randomized controlled trial of community-based dancing to modify disease progression in Parkinson disease. Neurorehabil Neural Repair 2012 Feb;26(2):132-143. [doi: 10.1177/1545968311421614] [Medline: 21959675]
5. Hackney ME, Earhart GM. Effects of dance on movement control in Parkinson's disease: a comparison of Argentine tango and American ballroom. J Rehabil Med 2010;42(7):662-669. [doi: 10.2340/16501977-0362]
6. Mirelman A, Rochester L, Reelick M, et al. V-TIME: a treadmill training program augmented by virtual reality to decrease fall risk in older adults: study design of a randomized controlled trial. BMC Neurol 2011;11(1):115. [doi: 10.1186/1471-2377-13-15]
7. Barry G, Galna B, Rochester L, Godfrey A. Delivery of community-based exercise programs for people with Parkinson's disease: a mixed-methods systematic review. J Neurol 2014;261(7):1321-1334.
8. Bevilacqua R, Benadduci M, Bonfigli AR, et al. Dancing with Parkinson's disease: the SI-ROBOTICS study protocol. Front Public Health 2021;9:780098. [doi: 10.3389/fpubh.2021.780098] [Medline: 34993171]
9. Cortellessa G, De Benedictis R, Fracasso F, Orlandini A, Umbrico A, Cesta A. AI and robotics to help older adults: revisiting projects in search of lessons learned. Paladyn J Behav Robot 2021 Feb 10;12(1):356-378. [doi: 10.1515/pjbr-2021-0025]
10. Carapellotti AM, Stevenson R, Doumas M. The efficacy of dance for improving motor impairments, non-motor symptoms, and quality of life in Parkinson's disease: a systematic review and meta-analysis. PLoS ONE 2020;15(8):e0236820. [doi: 10.1371/journal.pone.0236820] [Medline: 32756578]

11. Volpe D, Signorini M, Marchetto A, Lynch T, Morris ME. A comparison of Irish set dancing and exercises for people with Parkinson's disease: a phase II feasibility study. BMC Geriatr 2013 Jun 4;13:54. [doi: 10.1186/1471-2318-13-54] [Medline: 23731986]
12. Holden MK, Gill KM, Magliozzi MR. Gait assessment for neurologically impaired patients. Standards for outcome assessment. Phys Ther 1986 Oct;66(10):1530-1539. [doi: 10.1093/ptj/66.10.1530] [Medline: 3763704]
13. Holden MK, Gill KM, Magliozzi MR. Gait assessment for neurologically impaired patients. Phys Ther 1986 Oct 1;66(10):1530-1539. [doi: 10.1093/ptj/66.10.1530]
14. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988 May;19(5):604-607. [doi: 10.1161/01.str.19.5.604] [Medline: 3363593]
15. Rinaldi P, Mecocci P, Benedetti C, et al. Validation of the five-item geriatric depression scale in elderly subjects in three different settings. J Am Geriatr Soc 2003 May;51(5):694-698. [doi: 10.1034/j.1600-0579.2003.00216.x] [Medline: 12752847]
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
17. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
18. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986 Feb;34(2):119-126. [doi: 10.1111/j.1532-5415.1986.tb05480.x] [Medline: 3944402]
19. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
20. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
21. Balke B. A simple field test for the assessment of physical fitness. Rep 63-6. Rep Civ Aeromed Res Inst. 1963. URL: https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/1960s/media/am63-06.pdf
22. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
23. Ruggiero C, Mariani T, Gugliotta R, et al. Validation of the Italian version of the falls efficacy scale international (FES-I) and the SHORT FES-I in community-dwelling older persons. Arch Gerontol Geriatr 2009;49 Suppl 1:211-219. [doi: 10.1016/j.archger.2009.09.031] [Medline: 19836635]
24. Goetz CG, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): process, format, and clinimetric testing plan. Mov Disord 2007 Jan;22(1):41-47. [doi: 10.1002/mds.21198] [Medline: 17115387]
25. Umbrico A, Benadduci M, Bevilacqua R, et al. Combining clinical and spatial constraints into temporal planning to personalize physical rehabilitation. ICAPS 2023;33(1):532-540. [doi: 10.1609/icaps.v33i1.27234]
26. Belić M, Bobić V, Badža M, Šolaja N, Đurić-Jovičić M, Kostić VS. Artificial intelligence for assisting diagnostics and assessment of Parkinson's disease—a review. Clin Neurol Neurosurg 2019 Sep;184:105442. [doi: 10.1016/j.clineuro.2019.105442] [Medline: 31351213]
27. Victorino JN, Shibata Y, Inoue S, Shibata T. Predicting wearing-off of Parkinson's disease patients using a wrist-worn fitness tracker and a smartphone: a case study. Appl Sci (Basel) 2021;11(16):7354. [doi: 10.3390/app11167354]
28. Devarajan JP, Sreedharan VR, Narayanamurthy G. Decision making in health care diagnosis: evidence from Parkinson's Disease via hybrid machine learning. IEEE Trans Eng Manage 2023;70(8):2719-2731. [doi: 10.1109/TEM.2021.3096862]
29. Parisi L, RaviChandran N, Manaog ML. Feature-driven machine learning to improve early diagnosis of Parkinson's disease. Expert Syst Appl 2018 Nov;110:182-190. [doi: 10.1016/j.eswa.2018.06.003]
30. González JC, Pulido JC, Fernández F. A three-layer planning architecture for the autonomous control of rehabilitation therapies based on social robots. Cogn Syst Res 2017 Jun;43:232-249. [doi: 10.1016/j.cogsys.2016.09.003]
31. Baschieri D, Gaspari M, Zini F. A planning-based serious game for cognitive rehabilitation in multiple sclerosis. Presented at: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good; Nov 28, 2018; Bologna Italy. [doi: 10.1145/3284869.3284916]
32. Leone A, Rescio G, Diraco G, Manni A, Siciliano P, Caroppo A. Ambient and wearable sensor technologies for energy expenditure quantification of ageing adults. Sensors (Basel) 2022 Jun 29;22(13):4893. [doi: 10.3390/s22134893] [Medline: 35808387]
33. Co-Robotics. URL: http://www.corobotics.eu [accessed 2024-12-23]
34. Rajan K, Saffiotti A. Towards a science of integrated AI and robotics. Artif Intell 2017 Jun;247:1-9. [doi: 10.1016/j.artint.2017.03.003]
35. Ingrand F, Ghallab M. Deliberation for autonomous robots: a survey. Artif Intell 2017 Jun;247:10-44. [doi: 10.1016/j.artint.2014.11.003]
36. Umbrico A, De Benedictis R, Fracasso F, Cesta A, Orlandini A, Cortellessa G. A mind-inspired architecture for adaptive HRI. Int J Soc Robot 2023;15(3):371-391. [doi: 10.1007/s12369-022-00897-8] [Medline: 35910297]
37. Beale R, Creed C. Affective interaction: how emotional agents affect users. Int J Hum Comput Stud 2009 Sep;67(9):755-776. [doi: 10.1016/j.ijhcs.2009.05.001]

38. Umbrico A, Cesta A, Cortellessa G, Orlandini A. A holistic approach to behavior adaptation for socially assistive robots. Int J Soc Robotics 2020 Jul;12(3):617-637. [doi: 10.1007/s12369-019-00617-9]
39. ROS melodic morenia. ROS.org. URL: http://wiki.ros.org/melodic [accessed 2024-12-23]
40. Pstlab/roxanne_rosjava. GitHub. URL: https://github.com/pstlab/roxanne_rosjava.git [accessed 2024-12-02]
41. Umbrico A, Cesta A, Orlandini A. Human-aware goal-oriented autonomy through ROS-integrated timeline-based planning and execution. Presented at: The 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN); Aug 28-31, 2023; Busan, Republic of Korea. [doi: 10.1109/RO-MAN57019.2023.10309516]

Abbreviations

AI: artificial intelligence
BI: Barthel Index
FES-I: Falls Efficacy Scale-International
PD: Parkinson disease
POMA: Tinetti's Performance-Oriented Mobility Assessment
ROS: Robot Operating System
SF-12: 12-Item Short-Form Health Survey
SI-Robotics: Social Robotics for Active and Healthy Ageing
SPPB: Short Physical Performance Battery
TUG: Timed Up and Go
UPDRS-III : Unified Parkinson Disease Rating Scale – III

Edited by PH Chau; submitted 05.06.24; peer-reviewed by D Marks, F Noveletto, G Barresi; revised version received 05.09.24; accepted 09.09.24; published 14.01.25.

Please cite as:
Bevilacqua R, Maranesi E, Benadduci M, Cortellessa G, Umbrico A, Fracasso F, Melone G, Margaritini A, La Forgia A, Di Bitonto P, Potenza A, Fiorini L, La Viola C, Cavallo F, Leone A, Caroppo A, Rescio G, Marzorati M, Cesta A, Pelliccioni G, Riccardi GR, Rossi L
Exploring Dance as a Therapeutic Approach for Parkinson Disease Through the Social Robotics for Active and Healthy Ageing (SI-Robotics): Results From a Technical Feasibility Study
JMIR Aging 2025;8:e62930
URL: https://aging.jmir.org/2025/1/e62930
doi:10.2196/62930

Development and Validation of a Predictive Model Based on Serum Silent Information Regulator 6 Levels in Chinese Older Adult Patients: Cross-Sectional Descriptive Study

Yuzi You[1], MD; Wei Liang[2*], PhD; Yajie Zhao[2*], PhD

[1]Department of General Practice, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Wei Liang, PhD
Department of Geriatrics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai, China

Abstract

Background: Serum levels of silent information regulator 6 (SIRT6), a key biomarker of aging, were identified as a predictor of coronary artery disease (CAD), but whether SIRT6 can distinguish severity of coronary artery lesions in older adult patients is unknown.

Objectives: This study developed a nomogram to demonstrate the functionality of SIRT6 in assessing severity of coronary artery atherosclerosis.

Methods: Patients aged 60 years and older with angina pectoris were screened for this single-center clinical study between October 1, 2022, and March 31, 2023. Serum specimens of eligible patients were collected for SIRT6 detection by enzyme-linked immunosorbent assay. Clinical data and putative predictors, including 29 physiological characteristics, biochemical parameters, carotid artery ultrasonographic results, and complete coronary angiography findings, were evaluated, with CAD diagnosis as the primary outcome. The nomogram was derived from the Extreme Gradient Boosting (XGBoost) model, with logistic regression for variable selection. Model performance was assessed by examining discrimination, calibration, and clinical use separately. A 10-fold cross-validation technique was used to compare all models. The models' performance was further evaluated on the internal validation set to ensure that the obtained results were not due to overoptimization.

Results: Eligible patients (n=222) were divided into 2 cohorts: the development cohort (n=178) and the validation cohort (n=44). Serum SIRT6 levels were identified as both an independent risk factor and a predictor for CAD in older adults. The area under the receiver operating characteristic curve (AUROC) was 0.725 (95% CI 0.653 - 0.797). The optimal cutoff value of SIRT6 for predicting CAD was 546.384 pg/mL. Predictors included in this nomogram were serum SIRT6 levels, triglyceride glucose (TyG) index, and apolipoprotein B. The model achieved an AUROC of 0.956 (95% CI 0.928 - 0.983) in the development cohort. Similarly, in the internal validation cohort, the AUROC was 0.913 (95% CI 0.828 - 0.999). All models demonstrated satisfactory calibration, with predicted outcomes closely aligning with actual results.

Conclusions: SIRT6 shows promise in predicting CAD, with enhanced predictive abilities when combined with the TyG index. In clinical settings, monitoring fluctuations in SIRT6 and TyG may offer valuable insights for early CAD detection. The nomogram for CAD outcome prediction in older adult patients with angina pectoris may aid in clinical trial design and personalized clinical decision-making, particularly in institutions where SIRT6 is being explored as a biomarker for aging or cardiovascular health.

(JMIR Aging 2025;8:e64374) doi:10.2196/64374

KEYWORDS

aging; coronary artery disease; nomogram; SIRT6; TyG index; silent information regulator 6; triglyceride glucose index

Introduction

The global burden of ischemic heart disease from 1990 to 2019 reached staggering figures, with an estimated 9.14 million deaths attributed to coronary artery disease (CAD) in 2019 alone, affecting approximately 197 million individuals worldwide [1]. China witnessed a notable surge in CAD-associated mortality during this period, accounting for 38.2% of the global rise. By 2017, CAD-related deaths in China had skyrocketed by 1.12 million, representing an astonishing 184.1% increase compared with that of 1990 [2,3]. As of 2019, China was the global leader in CAD-related deaths, with 1.87 million reported cases [4]. By 2029, a 64% increase in CAD cases is expected in China

compared with 2020. As a dynamically evolving condition, CAD significantly contributes to the development of major adverse cardiovascular events, such as myocardial infarction, stroke, and cardiovascular mortality [5].

Aging, an independent risk factor for CAD, is characterized by a progressive decline in coronary artery and microvascular function, resulting in altered myocardial perfusion and increased myocardial injury [6]. CAD predominantly affects middle-aged and older adult populations, particularly people aged 60 years and older, with a prevalence exceeding 70%. Notably, the older adult population in China, aged 65 years and older, reached 220 million in 2023, constituting 15.4% of the nation's population and approximately 26.8% of the global older adult population [7]. The aging demographic in China is expanding rapidly, with nearly 40 million individuals aged 80 years and older in 2022, accounting for approximately 2.7% of the total population of China. Hence, China is predicted to enter a super-aged society by 2030, with the older adult population comprising more than 20% of the total population [8]. By 2084, it is estimated that half of China's population will be older adults. A higher mortality rate has been observed among Chinese older adults with CAD, with the majority of deaths occurring in people aged 75 years and older [9]. Despite these trends, research on CAD has often overlooked the specific needs of the older adult population. In China, approximately 300 million people suffer from chronic diseases, with half of them being aged 65 years or older. The older adult population often presents with atypical clinical symptoms, multiple comorbidities, and prolonged use of medications, all of which amplify the risk of cardiovascular disease [10,11]. Previous studies have indicated that therapeutic interventions in the early stages of CAD can reduce its incidence and improve prognosis [12,13], underscoring the necessity for noninvasive methods for identifying predictive factors for CAD.

Silent information regulator 6 (SIRT6), a member of the nicotinamide adenine dinucleotide–dependent histone deacetylase family, plays a crucial role in aging and aging-related diseases by maintaining telomerase stability and metabolic homeostasis, as well as regulating oxidative stress [14]. Recent evidence has highlighted the protective properties of SIRT6 against CAD in preserving endothelial function, inhibiting inflammatory responses, and regulating glucose and lipid metabolism. Hence, a decline in serum SIRT6 levels has emerged as an independent risk factor for CAD [15], warranting further investigation of its diagnostic and predictive value in older adult patients. Existing risk assessment models, such as the Pooled Cohort Equations model and the Systematic Coronary Risk Estimation model, were primarily developed based on Caucasian and African American populations and lack specific biomarkers and demonstrate poor calibration when applied to the Chinese population [16]. Given these ethnic differences, there is an urgent need for tailored risk assessment models in cardiovascular disease prevention, particularly in diverse populations such as China.

In this study, we aimed to develop a nomogram for predicting outcomes in older adults presenting with clinical symptoms suggestive of CAD. We hypothesized that a combination of baseline SIRT6 levels with clinical parameters could improve the evidence-based selection of candidates for this marker and facilitate clinical decision-making, resulting in its potential implementation in clinical trials.

Methods

Study Design and Participants

In this single-center study, a nomogram was developed for predicting the outcomes of potential cases with CAD and was validated using data from Ruijin Hospital, Shanghai Jiao Tong University School Of Medicine. Patients aged 60 years or older who were diagnosed with angina pectoris were screened between October 1, 2022, and March 31, 2023. The Judkins method [17] was used for performing coronary angiography (CAG) via the radial or femoral artery. The angiographic results underwent joint assessment by 3 experienced cardiovascular specialists (WFS , ZBZ, and JWN). The severity of stenosis in the major coronary arteries, including the left anterior descending branch (LAD), left circumflex artery, right coronary artery, and left main coronary artery, was evaluated, with CAD defined as ≥50% stenosis in any 1 vessel and coronary atherosclerosis (CAS) as <50% stenosis in all vessels.

Eligible participants were individuals aged 60 years or older who presented with chest pain and had been evaluated by a specialist, resulting in a preliminary diagnosis of CAD with an indication for CAG. Participants must have no prior history of CAG, percutaneous coronary intervention, or coronary artery bypass grafting. Furthermore, informed consent must be obtained for the collection of biological samples, with blood samples obtained prior to CAG. Participants must also have coronary artery stenosis as confirmed by the CAG results. Exclusion criteria included recent acute infections, gastrointestinal bleeding, surgical procedures, or trauma within the past 6 months. Individuals were also excluded if they had positive viral markers, including hepatitis B surface antigen, hepatitis B core antibody with hepatitis B virus–DNA of the detection threshold or greater, positive hepatitis C virus antibody with hepatitis C virus–RNA of the upper limit of normal or greater, or positive HIV antibody. Severe cardiac conditions also warranted exclusion, including decompensated heart failure, significant valvular heart disease within the past 6 months, notable electrocardiographic abnormalities (eg, any degree of atrial fibrillation, second-degree type II or third-degree atrioventricular block, or corrected QT interval exceeding 470 milliseconds in females or 450 milliseconds in males), uncontrolled symptomatic arrhythmias, cerebrovascular events or transient ischemic attacks within the past 6 months, a history of malignancy or autoimmune diseases, or severe liver or renal disorders unrelated to the study condition.

Upon admission, patients were assessed for CAD severity using the Gensini score, an angiographic tool for grading coronary artery lesions [18]. Serum specimens were collected from eligible patients for the detection of SIRT6 levels by enzyme-linked immunosorbent assay. Clinical data and the complete CAG inspection results of these participants were also collected.

Ethical Considerations

The study protocol was approved by the human ethics committee of Ruijin Hospital (KY2021-108 Ruijin Hospital). It adhered to strict data confidentiality measures in compliance with the Helsinki Declaration and the institutional guidelines and reporting studies conducted using routinely collected observational data. Participants provided written informed consent at the time of data collection.

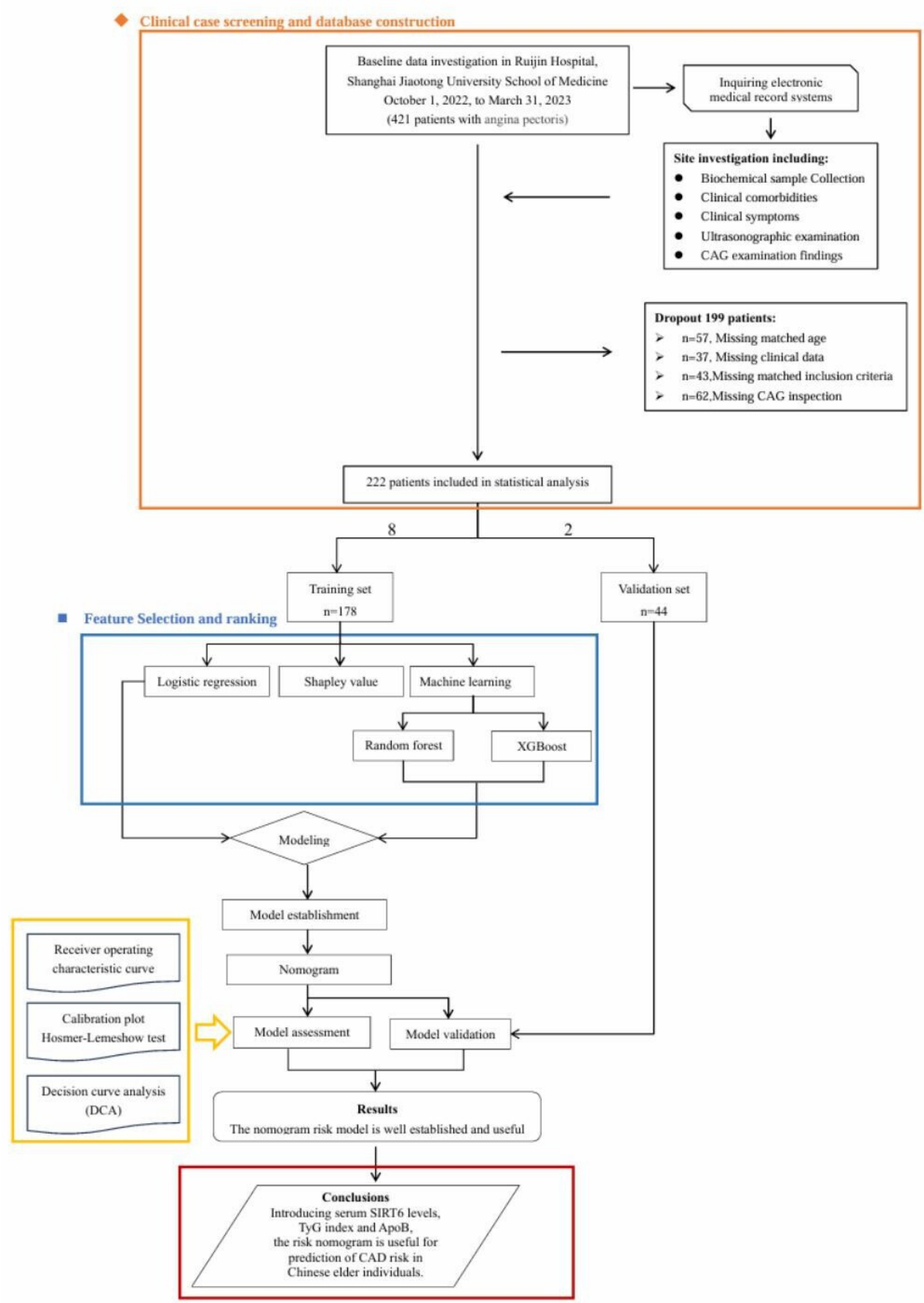
Selection Bias

Several rigorous measures were implemented to address potential selection bias in this study. Stringent inclusion and exclusion criteria were carefully defined and applied to ensure a well-characterized and homogeneous study population. Furthermore, random sampling techniques were used where appropriate, and blinding in outcome assessment was ensured to minimize bias in both participant selection and data interpretation. These strategies were designed to enhance the internal validity and robustness of our study results.

Data Collection

All recruited patients were divided into 2 cohorts, the development cohort and the validation cohort, in an approximate ratio of 8:2. Then, predefined criteria were applied to ensure cohort comparability. A total of 29 pretherapeutic parameters were collected, including demographic characteristics, initial symptoms, history of hypertension and diabetes mellitus, baseline clinical status, and baseline laboratory test results. Finally, 3 out of 26 collected parameters were tested as putative predictors for outcomes in the models. Three putative predictors allowed for 9-10 events per predictor for the primary outcome in the training cohort, meeting the recommended minimal number of events per predictor. These putative predictors were selected based on previous research demonstrating their potential prognostic value in CAD and the investigators' (WFS, ZYC, and JWN) clinical experience with CAD. The study flowchart is shown in Figure 1.

Figure 1. Study flowchart. ApoB: apolipoprotein B; CAD: coronary artery disease; CAG:coronary angiography; SIRT6: silent information regulator 6; TyG: triglyceride glucose; XGBoost: Extreme Gradient Boosting.



Statistical Analysis

Variable Selection

In this study, IBM SPSS Statistics (version 26.0; IBM Corp) software was used for baseline description and logistic regression (LR) analysis. Continuous variables were presented as mean (SD), while categorical variables were expressed as percentages (n [%]). Normality was assessed using the Shapiro-Wilk test, with $P>.05$ indicating adherence to normal distribution. For normally distributed variables, an intergroup analysis was conducted using the *t* test, while nonnormally distributed data were compared using the Wilcoxon rank sum test. Features with >30% missing values were removed. The remaining 29 features were collected for further processing. For features with missing values of <10%, median or mean imputation was used. Those with 10% - 20% missing values were imputed by the "MissForest" package in R (R Foundation for Statistical Computing). Random forest and Extreme Gradient Boosting (XGBoost) algorithms were implemented via R statistical software (version 4.3.2; R Foundation for Statistical Computing). Receiver operating characteristic (ROC) curves were plotted using the "pROC" package, while nomograms were constructed using the "rms" and "regplot" packages. The "val.prob" function was used for refining calibration curve plots and Hosmer-Lemeshow tests, while the "dcurves" function facilitated decision curve analysis (DCA). All statistical tests were 2-tailed, with $P<.05$ considered statistically significant. Feature ranking was obtained using Shapley Additive Explanation, Gini, and Gain values, respectively. A 10-fold cross-validation strategy was applied to develop the data set into training, validation, and test sets.

Model Building and Visualization:

Three methodologies, including LR, random forest, and XGBoost, were used for model building. All analyses were conducted using the statistical software package R (version 4.3.2). The data from the Ruijin Hospital of Shanghai Jiao Tong Medical University were used as the training set and internal validation set for model development and verification. Binary LR, random forest, and XGBoost models were constructed using the training data to classify patients into those with CAS and those with CAD. The XGBoost model exhibited the highest comprehensive discriminant ability [19] and therefore was selected for further analysis. The final model was visualized as a column line chart to address issues of poor machine learning interpretability and consequent low clinical use.

Model Comparison

Accuracy, precision, recall, and F_1-score were used to evaluate the performance of the multiclassification model. The model error was further analyzed using a confusion matrix. Discrimination, calibration, and clinical use between the nomogram and the variables incorporated into the nomogram were assessed in both the training and validation sets, respectively. ROC curves, calibration curves, and DCA plots were plotted, and the areas under the ROC curves (AUROCs) were compared using the Delong test. Calibration was evaluated using the Hosmer-Lemeshow test.

Results

Characteristics of the Training and Validation Sets

The final study cohort comprised 222 patients, with 178 allocated to the training set and 44 to the validation set. The demographic characteristics and clinical results of patients in both sets are summarized in Table 1. There were 23.87% (53/222) of patients in the CAS group and 76.13% (169/222) of patients in the CAD group. There were significant differences in the low-density lipoprotein cholesterol (LDL-C), apolipoprotein B (ApoB), apolipoprotein E (ApoE), fasting blood glucose (FBG), hemoglobin A_{1c} (HbA_{1c}), Gensini score, TyG index, and atherogenic index of plasma (AIP) between the CAS and CAD groups ($P<.05$). Specifically, the CAD group exhibited a higher TyG index (12.4 [SD 1.38]) and AIP (0.10 [SD 0.31]) than the CAS group (10.9 [SD 0.54], 0.00 [SD 0.27]), as shown in Table 1. Significant differences were also observed in categorical variables, such as sex, history of diabetes, clinical symptoms, segmental vascular lesions (ie, LAD, left circumflex artery, left main coronary artery, and right coronary artery), and 10-year cardiovascular risk between the CAS and CAD groups ($P<.05$).

Table . Baseline characteristics of participants[a].

Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
Age (years), mean (SD)	69.3 (5.33)	68.9 (5.12)	.56
Sex, n (%)			.03*
Female	28 (52.8)	59 (34.9)	
Male	25 (47.2)	110 (65.1)	
BMI, mean (SD)	25.1 (4.46)	24.6 (2.64)	.50
Drinking, n (%)			1.00
No	47 (88.7)	151 (89.3)	
Yes	6 (11.3)	18 (10.7)	
Smoking, n (%)			.68
No	47 (88.7)	144 (85.2)	
Yes	6 (11.3)	25 (14.8)	
SBP[d], mean (SD)	145 (18.3)	146 (17.8)	.82
DBP[e], mean (SD)	77.6 (10.1)	78.5 (10.8)	.60
Diabetes mellitus, n (%)			.006**
No	43 (81.1)	100 (59.2)	
Yes	10 (18.9)	69 (40.8)	
Hypertension, n (%)			.65
No	20 (37.7)	56 (33.1)	
Yes	33 (62.3)	113 (66.9)	
hs-CRP[f], mean (SD)	2.19 (3.37)	2.15 (3.32)	.94
CysC[g], mean (SD)	1.00 (0.16)	1.59 (7.08)	.28
eGFR[h], mean (SD)	84.1 (9.63)	82.9 (10.6)	.47
mALB[i], mean (SD)	1.40 (1.78)	2.03 (5.04)	.17
cTnI[j], mean (SD)	7.60 (16.1)	25.3 (114)	.05
proBNP[k], mean (SD)	170 (338)	305 (1764)	.35
LDL-C[l], mean (SD)	2.36 (0.89)	1.95 (0.69)	.003**
Lp(a)[m], mean (SD)	0.26 (0.32)	0.31 (0.35)	.35
sLDL[n], mean (SD)	0.76 (0.35)	0.66 (0.31)	.08
ApoA1[o], mean (SD)	1.35 (0.25)	1.28 (0.24)	.10
ApoB[p], mean (SD)	0.76 (0.23)	0.68 (0.19)	.02*
ApoE[q], mean (SD)	3.77 (0.52)	3.60 (0.62)	.05
FBG[r], mean (SD)	5.67 (1.00)	6.10 (1.72)	.03*
HbA_{1c}[s], mean (SD)	6.00 (0.66)	6.46 (1.28)	.001**
FINS[t], mean (SD)	9.87 (5.37)	11.7 (11.3)	.11
CP[u], mean (SD)	2.61 (1.16)	2.72 (1.23)	.56
Carotid plaque, n (%)			.049*
<50%	37 (69.8)	138 (81.7)	
>50%	1 (1.89)	9 (5.33)	

Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
No	15 (28.3)	22 (13)	
Ten-year cardiovascular risk, n (%)			.009**
High	21 (39.6)	107 (63.3)	
Low	7 (13.2)	15 (8.88)	
Middle	25 (47.2)	47 (27.8)	
LAD [v] , n (%)			.01*
No	11 (20.8)	12 (7.10)	
Yes	42 (79.2)	157 (92.9)	
RCA [w] , n (%)			<.001***
No	24 (45.3)	27 (16)	
Yes	29 (54.7)	142 (84)	
LCX [x] , n (%)			<.001***
No	38 (71.7)	57 (33.7)	
Yes	15 (28.3)	112 (66.3)	
LM [y] , n (%)			.01*
No	53 (100)	147 (87)	
Yes	N/A[z]	22 (13)	
SIRT6[aa], mean (SD)	866 (510)	495 (443)	<.001***
TyG[ab], mean (SD)	10.9 (0.54)	12.4 (1.38)	<.001***

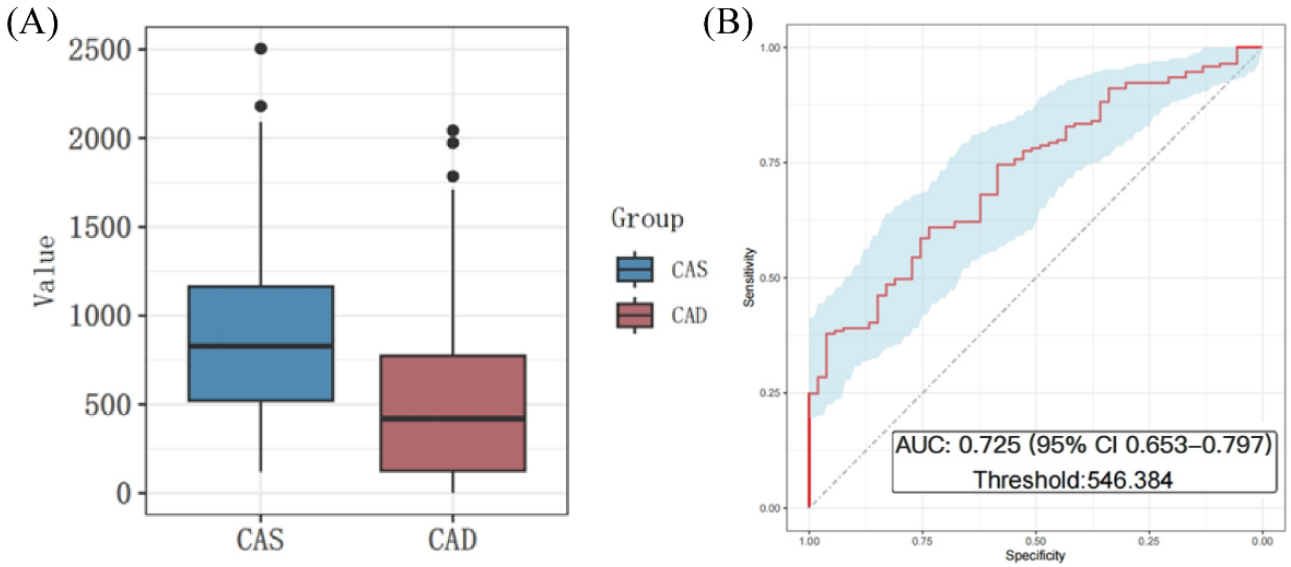
Variables	CAS[b] (n=53)	CAD[c] (n=169)	*P* value
AIP[ac], mean (SD)	0.00 (0.27)	0.10 (0.31)	.02*

[a]Variables of significance (*$P \leq .05$, ** $P \leq .01$, and *** $P \leq .001$).
[b]CAS, coronary atherosclerosis.
[c]CAD: coronary artery disease.
[d]SBP: systolic pressure.
[e]DBP: diastolic blood pressure.
[f]hs-CRP: hypersensitive C-reactive protein.
[g]CysC: cystatin c.
[h]eGFR: estimated glomerular filtration rate.
[i]mALB: microalbumin.
[j]cTnI: troponin I.
[k]proBNP: pro-B-type natriuretic peptide.
[l]LDL-C: low-density lipoprotein cholesterol.
[m]Lp(a): lipoprotein(a).
[n]sLDL: small dense low-density lipoprotein.
[o]ApoA1: apolipoprotein A1.
[p]ApoB: apolipoprotein B.
[q]ApoE: apolipoprotein E.
[r]FBG: fasting blood glucose.
[s]HbA_{1c}: hemoglobin A_{1c}.
[t]FINS: fasting insulin.
[u]CP: C-peptide.
[v]LAD: left anterior descending branch.
[w]RCA: right coronary artery.
[x]LCX: left circumflex artery.
[y]LM: left main coronary artery.
[z]N/A: not applicable.
[aa]SIRT6: silent information regulator 6.
[ab]TyG: triglyceride glucose.
[ac]AIP: atherogenic index of plasma.

Circulating SIRT6 Levels in Patients

To investigate the potential difference in serum SIRT6 levels between the CAD and CAS groups, the enzyme-linked immunosorbent assay was used to determine the SIRT6 levels in serum samples. A significant reduction in serum SIRT6 levels was observed in the CAD group (495 [SD 443] pg/mL) compared with the CAS group (866 [SD 510] pg/mL) (Figure 2A). To further assess the diagnostic potential of serum SIRT6 as a biomarker for distinguishing between CAS and CAD, ROC analysis was performed. A serum SIRT6 level of 546.384 pg/mL was identified as the optimal cutoff value for discriminating between CAS and CAD (AUROC 0.725, 95% CI 0.653 - 0.797) (Figure 2B).

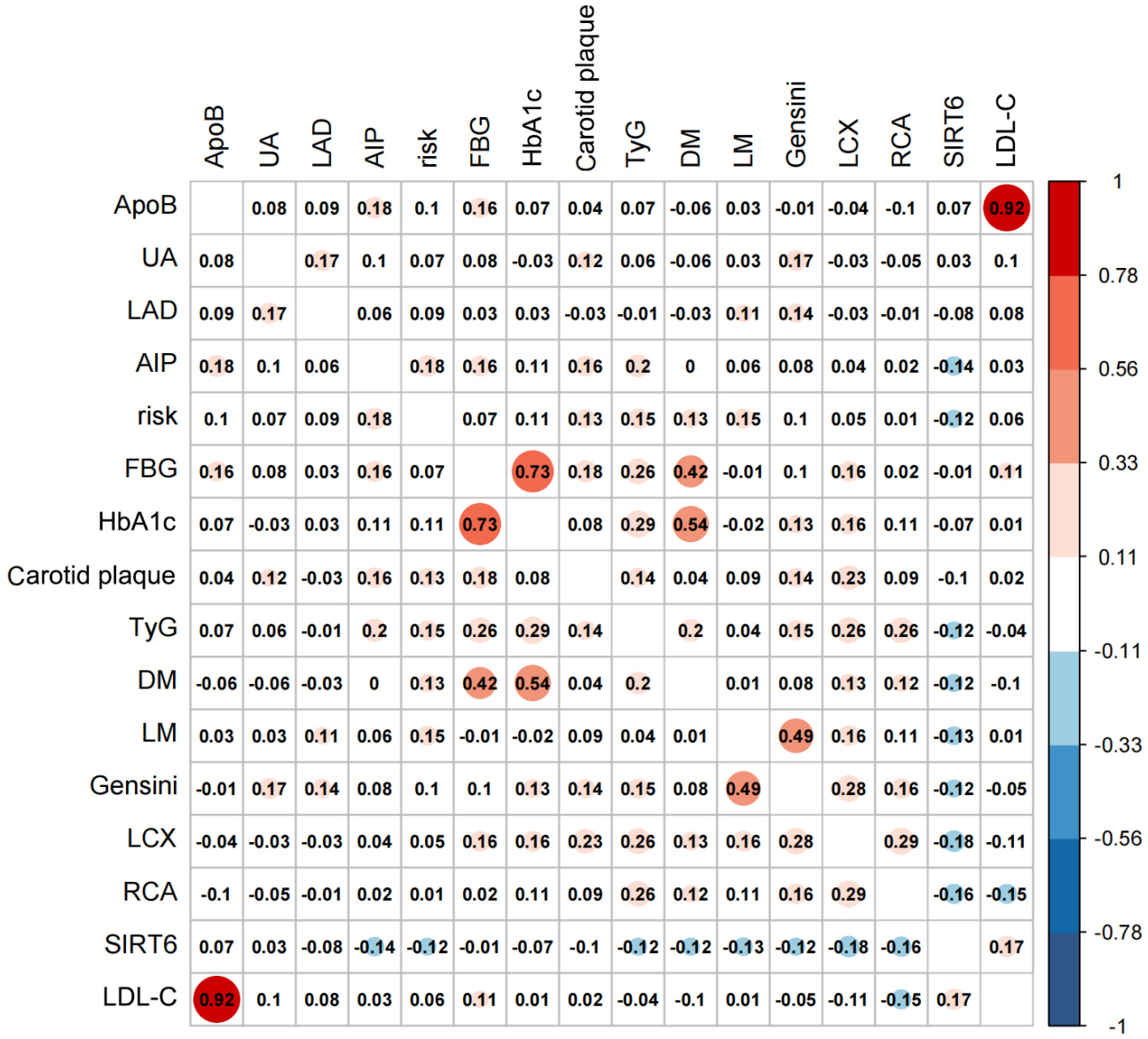
Figure 2. Serum levels of silent information regulator 6 (SIRT6) and receiver operating characteristic (ROC) curve. (**A**) Serum levels of SIRT6 were lower in patients with CAD and patients with CAS. (**B**) ROC curve analysis of SIRT6 in predicting the diagnosis of CAD (AUROC 0.725, 95% CI 0.653 - 0.797). AUROC: area under the receiver operating characteristic curve; CAD: coronary artery disease; CAS: coronary atherosclerosis.



Exploration of Variable Correlations

Factors with significant differences in intergroup analysis, including ApoB, unstable angina pectoris, LDL-C, AIP, FBG, HBA_{1c}, TyG index, history of diabetes mellitus, SIRT6, LAD, AIP, 10-year cardiovascular risk, carotid plaque burden, segmental vascular lesions, and Gensini score, were selected for further interfactor correlation analysis. To illustrate the correlations among these predictive indicators, a correlation coefficient matrix graph was constructed. While serum SIRT6 did not exhibit positive correlations with other indicators, aside from LDL-C, it demonstrated negative correlations with the Gensini score, 10-year cardiovascular risk, and carotid plaque burden (Figure 3).

Figure 3. Correlation heatmap analysis. AIP: atherogenic index of plasma; ApoB: apolipoprotein B; DM: diabetic mellites; FBG: fasting blood glucose; HbA_{1c}: hemoglobin A_{1c}; LAD: left anterior descending branch; LCX: left circumflex artery; LDL-C: low-density lipoprotein cholesterol; LM: left main coronary artery; RCA: right coronary artery; Risk: 10-year cardiovascular risk; SIRT6: silent information regulator 6; TyG: triglyceride glucose; UA: unstable angina pectoris.

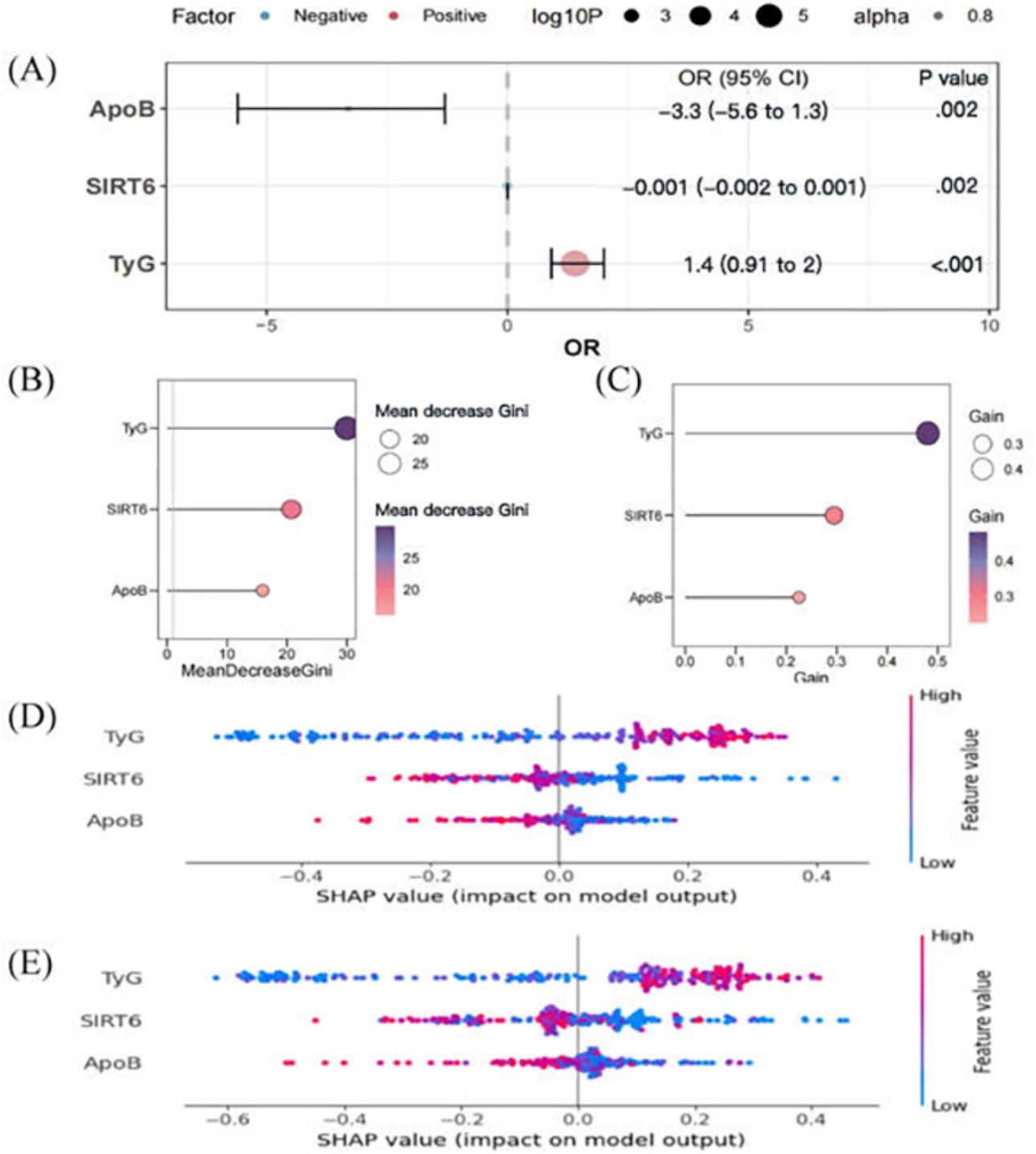


A correlation coefficient matrix graph was constructed. Darker shades represent stronger correlations, with positive correlations indicated by positive values and negative correlations by negative values. The predictive indicators included in this analysis are primarily independent predictive factors.

Feature Selection and Ranking

The significance of variables was preliminarily evaluated by LR analysis. Ultimately, serum SIRT6 levels, TyG index, and ApoB were incorporated into the model construction (Figure 4A). Feature importance ranking was determined using Gini (Figure 4B), Gain (Figure 4C), and Shapley Additive Explanation values (Figure 4D and E), which yielded similar results (Figure 4).

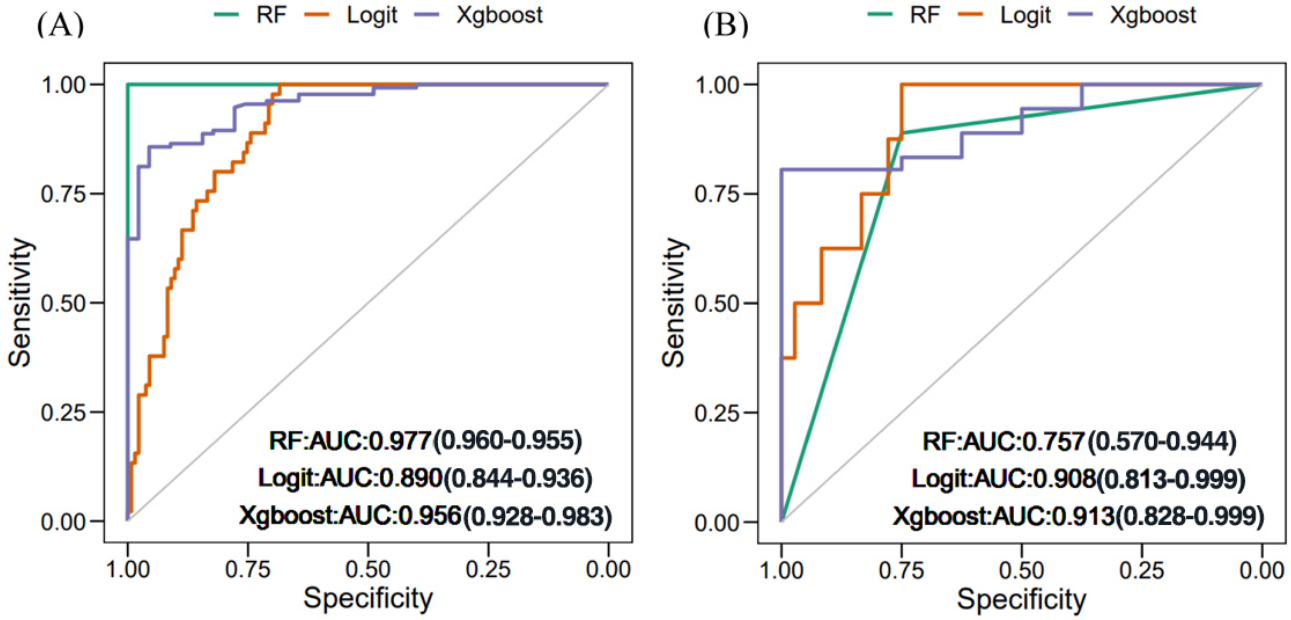
Figure 4. Feature importance. (**A**) Logistic regression results; (**B**) random forest results; (**C**) XGBoost results; (**D**) random forest SHAP; and (**E**) XGBoost SHAP. ApoB: apolipoprotein B; OR: odds ratio; SHAP: Shapley Additive Explanation; SIRT6: silent information regulator 6; TyG: triglyceride glucose; Extreme Gradient Boosting XGBoost.



Model Performance Comparisons

An LR model and 2 machine learning models were constructed to predict the development of CAD in older adult patients. The discriminative performance of the 3 models is shown via ROC curves in Figure 5.

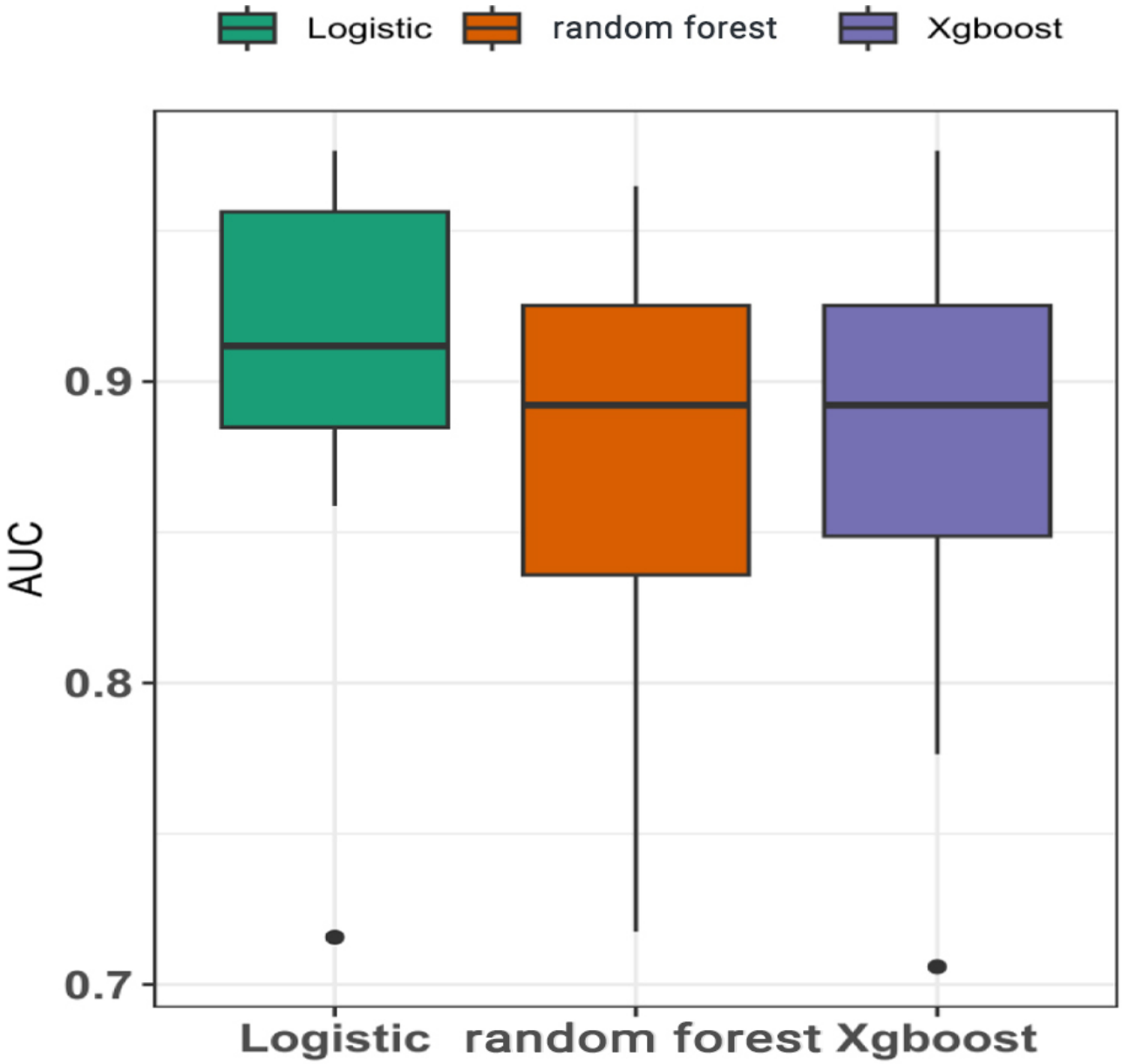
Figure 5. Receiver operating characteristic curves for predicting different classes using various models. (**A**) AUROC of the training set and (**B**) AUROC of the validation set. AUROC: area under the receiver operating characteristic curve; RF: random forest; Xgboost: Extreme Gradient Boosting.



In the training set, the random forest model demonstrated the best predictive ability for CAD in older adult patients (AUROC 0.977, 95% CI 0.960 - 0.995), followed by the XGBoost (AUROC 0.956) and LR (AUROC 0.890) models (Figure 5A). However, in the validation set, the classification performance of the random forest model was noticeably inferior to that of the XGBoost model. The AUROC values of XGBoost, LR, and random forest models were 0.913, 0.908, and 0.757, respectively (Figure 5B).

In cross-validation, the LR model demonstrated the best predictive ability for CAD in older adult patients (AUROC 0.912), followed by the XGBoost (AUROC 0.892) and random forest (AUROC 0.892) models (Figure 6).

Figure 6. Cross-validation AUROC of various models. AUROC: area under the receiver operating characteristic curve; Xgboost: Extreme Gradient Boosting.



Detailed performance metrics of the 3 models are shown in Table 2. The XGBoost model exhibited the best discrimination, with the highest recall (1) and accuracy (0.863), the second-highest precision (0.833), and the second-highest F_1-score (0.909). Although the XGBoost model may not perform as well as the other 2 models in certain aspects, its overall stability and balanced performance on both the training and the validation sets make it a more reliable choice. As a results, we developed a CAD prediction model for the older adults using the XGBoost algorithm and visualized the results through a nomogram. The nomogram based on the serum SIRT6-level model is shown in Figure 7.

Table . Model performance metrics.

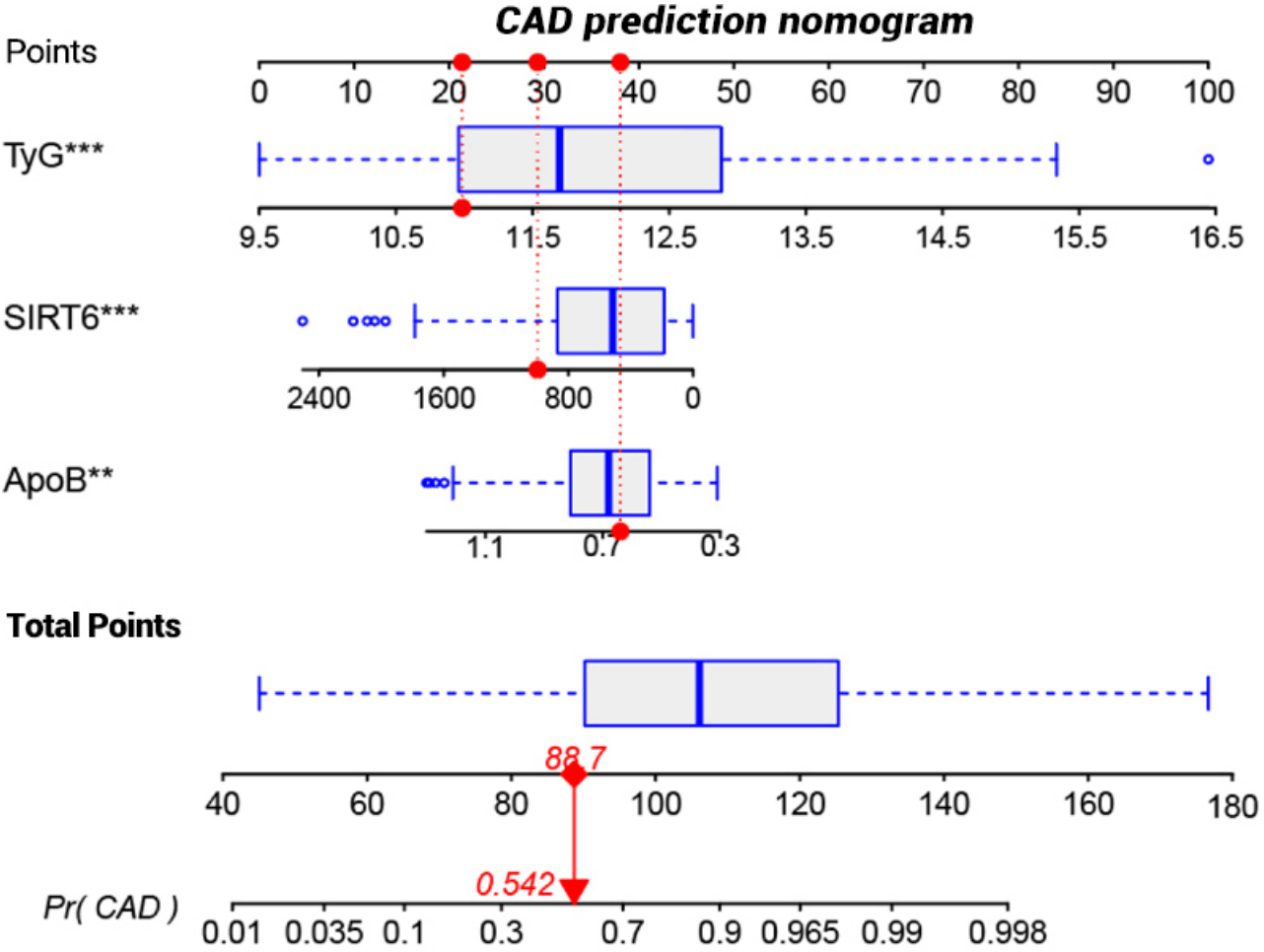
Model	Accuracy	Precision	Recall	F_1-score
LR[a]	0.795	0.75	1	0.857
RF[b]	0.863	0.941	0.889	0.914
XGBoost[c]	0.863	0.833	1	0.909

[a]LR: logistic regression.

[b]RF: random forest.

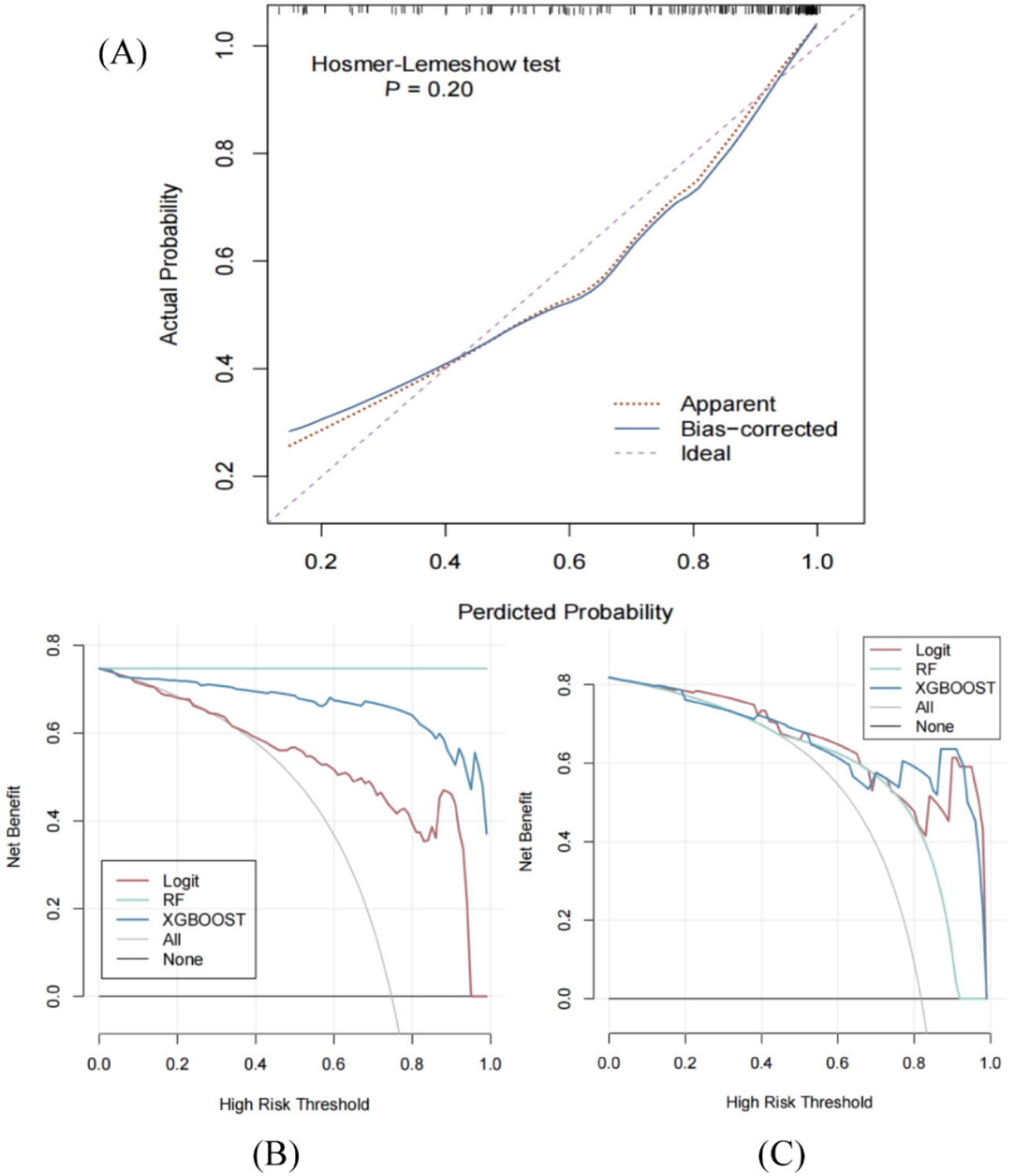
[c]XGBoost: Extreme Gradient Boosting.

Figure 7. Nomogram for predicting the probability of CAD. Variables of significance (**P≤.01 and ***P≤.001). ApoB: apolipoprotein B; CAD: coronary artery disease; SIRT6: silent information regulator 6; TyG: triglyceride glucose.



The calibration curves (Figure 8A) showed a similar trend among the 3 models, and the Hosmer-Lemeshow test results (χ^{28}=11.001; P=.20) indicated no significant difference between the predicted and observed values. These data suggest that the XGBoost model has good calibration ability.

Figure 8. Calibration curve and decision curve analyses for predicting coronary artery disease. (**A**) Calibration curves of the training cohort; (**B**) decision curve analysis (DCA) of the training cohort; and (**C**) DCA of the validation cohort. RF: random forest; XGBoost: Extreme Gradient Boosting.



The DCA revealed that the XGBoost model exhibited greater net benefit along with the threshold probability, indicating its superior clinical use compared with other models (Figure 8B and C).

Discussion

Principal Findings

An accurate risk assessment is critical for patient-centered clinical decision-making. The nomogram presented in this study may serve as an important clinical decision aid tool, assisting in determining whether invasive CAG should be performed, or informing patients as a basis for joint decision-making. In this study, traditional LR and 2 common machine learning

algorithms were used to establish a prediction and evaluation model of coronary artery stenosis in older adult patients with suspected CAD. Through a comprehensive analysis of models, the XGBoost algorithm–based model with excellent prediction performance and good internal verification ability was selected. In addition, we developed a nomogram for CAD prediction by combining traditional clinical variables (eg, ApoB) with 2 novel variables relevant to the older adult population, serum SIRT6 levels and the TyG index, which allowed for assessing the risk of coronary artery stenosis more conveniently and interactively (Figure 6). This, to some extent, helps address the issue of poor interpretability, which often hampers the implementation of machine learning models.

Given its significant role in genomic stability, DNA repair, and telomere function, SIRT6 has been considered to have therapeutic potential in aging and aging-related diseases. Moreover, SIRT6 plays a crucial role in the development and progression of CAD due to its involvement in oxidative stress, inflammation, and energy metabolism [20]. Previous studies have highlighted the importance of SIRT6 in protecting blood vessels and the heart from endothelial dysfunction, atherosclerosis, myocardial fibrosis, and ischemia or reperfusion injury [21,22].

In this study, we observed lower serum SIRT6 levels in older adult patients with CAD than in those with CAS. Correlation analysis revealed a negative association between serum SIRT6 levels and the Gensini score, suggesting that reduced serum SIRT6 levels may exacerbate the severity of coronary artery lesions. Furthermore, serum SIRT6 levels were negatively correlated with segmental vascular lesions and carotid plaque burden, indicating that serum SIRT6 may reflect the overall burden of atherosclerotic plaques in blood vessels. Dyslipidemia is an independent risk factor for the occurrence and development of CAD, while diabetes mellitus doubles the risk of cardiovascular disease and is a leading cause of mortality in patients with dyslipidemia [23,24].

LDL-C levels are considered a determinant of the absolute risk of major cardiovascular events [25]. Previous evidence has revealed that liver-specific knockout of SIRT6 significantly increases the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) gene expression and plasma LDL-C levels, while SIRT6 overexpression improves lipid metabolism and reduces atherosclerosis risk by lowering plasma LDL-C and *PCSK9* levels [26]. Although a previous clinical study in a Chinese population of all ages found no correlation between serum SIRT6 and LDL-C levels [15], in this study, we found a positive correlation between serum SIRT6 and LDL-C levels. Future investigations are needed to explore the mechanisms by which SIRT6 affects cardiovascular function through lipid metabolism in the older adult population.

Disruption of glucose homeostasis is another recognized risk factor for CAD. Animal models have shown that SIRT6 plays an important role in glucose production and uptake, insulin signaling, and metabolism [27]. Our data revealed that circulating SIRT6 levels were negatively correlated with FBG and the comorbidity of diabetes, which is consistent with previous findings [28]. Whether glucose metabolism disruption mediates the relationship between SIRT6 and CAD warrants further investigation. In addition, our study identified serum SIRT6 as an independent biomarker for assessing the severity of coronary artery lesions in older adults, with the optimal cutoff value for classifying atherosclerosis and CAD being 546.384 pg/mL.

Abnormal blood lipids are established independent risk factors for CAD. Unlike previous studies that commonly included LDL-C, HDL-C, and other conventional measures, our study incorporated ApoB as one of the predictive factors. ApoB is a crucial component of LDL-C, and its deposition within the arterial wall is a fundamental step driving the progression of atherosclerosis, from initial lipid deposition to the development of acute complex events, such as plaque rupture [29]. Compared with other individual lipid markers, ApoB serves as a more accurate biomarker for cardiovascular risk assessment [30,31]. Elevated ApoB levels, representing LDL levels, are associated with the occurrence of CAD, and ApoB levels are positively correlated with the Gensini score, indicating the degree of atherosclerosis and arterial narrowing [32]. In addition, circulating ApoB levels are considered more predictive than plasma cholesterol levels for early-onset CAD risk [33]. Correlation analysis indicated a significant positive correlation between ApoB and LDL-C levels. Moreover, in all 3 statistical methods used for feature selection, ApoB consistently emerged as significant, which is in line with previous findings [31]. However, in an intergroup analysis, we observed higher levels of ApoB in the CAS group than in the CAD group, which is contrary to previous results [33].

Insulin resistance further disturbs glucose and lipid metabolism in patients with diabetes, promoting chronic inflammation, disrupting normal endothelial function, and accelerating the development of complications, such as atherosclerosis-related cardiovascular diseases. The persistent prevalence of insulin resistance is considered a major contributor to the high mortality rate in atherosclerosis-related cardiovascular diseases worldwide [34]. The TyG index is a quantitative measure based on blood glucose and triglycerides that assesses insulin resistance. A clinical study has shown an association between the TyG index and coronary artery calcification [35]. However, in cohort studies, the TyG index was found to be correlated with carotid atherosclerosis [36] but unrelated to the incidence of CAD [37]. In our study, intergroup analysis revealed that the TyG index was significantly higher in the CAD group than in the CAS group. Correlation analysis indicated positive associations between the TyG index and the Gensini score, carotid plaque burden, and segmental vascular lesions. Furthermore, in multifactorial regression analysis and machine learning feature selection, the TyG index consistently demonstrated a strong correlation with CAD, suggesting that it may be an independent risk factor for CAD. Furthermore, we observed a negative correlation between serum SIRT6 levels and the TyG index. Future research is needed to explore the impact of SIRT6 on the TyG index and to elucidate the relationship between these factors in terms of glucose and lipid metabolism disturbance and the development of CAD.

Strengths and Limitations

Our prediction model addresses the challenge of poor interpretability associated with machine learning models. By incorporating readily obtainable clinical and physiological indicators, this nomogram not only provides prediction results and probabilities directly but also facilitates personalized intervention through probability curves. Tailored patient treatment may reduce medical costs and unnecessary invasive tests. Particularly in resource-constrained medical settings, our model may assist in disease screening and alleviate the medical burden.

Although the overall performance of the model was good, the sample distribution in our study was not balanced, and the model was not validated externally with independent data sets. Therefore, the generalizability and extrapolation to the overall population cannot be currently estimated. Second, the model was more inclined to classify patients without the disease as having the disease, while the proportion of patients with the disease misclassified as being disease-free was very low. The rate of missed diagnosis was also low. Third, the limited small sample size and single-center, cross-sectional design constrained the generalizability of the results. Furthermore, this study specifically focused on the older adult population; however, there is currently no consensus on age-related changes in circulating ApoB levels. Furthermore, some studies have suggested that the ApoB–apolipoprotein A1 (ApoA1) ratio has a higher predictive value for atherosclerosis and intima-media thickness than individual lipid markers. Therefore, in future research, expanding the sample size and using the ApoB-ApoA1 ratio as a composite indicator are needed to further evaluate the correlation between ApoB and the severity of coronary artery lesions.

Conclusions

Based on the LR, random forest, and XGBoost algorithms, we developed a predictive model for CAD occurrence in the older adult population, incorporating SIRT6, the TyG index, and ApoB. The model was comprehensively evaluated for discrimination, calibration, clinical applicability, and internal validation. The XGBoost predictive model exhibited favorable predictive performance and clinical use, which may facilitate early CAD screening and diagnosis in older adult patients, particularly for identifying high-risk individuals. Furthermore, the model may reduce unnecessary invasive examinations in negative patients and minimize missed diagnoses in positive patients. The personalized probability curve generated by the model may offer targeted intervention guidance. Our findings also underscore the importance of considering risk factors such as SIRT6, ApoB, and TyG levels in this population.

Acknowledgments

This research is supported by a grant from the National Natural Science Foundation of China (grant 82103079) and the Shanghai Municipal Health Commission Clinical Research in Health Industry Project (grant 202140052).

Authors' Contributions

YY conceptualized and designed the study, collected and analyzed data, drafted the manuscript, and implemented the study. WL and YZ provided the study materials instrumentation and other analysis tools and checked literature. All authors had full access to all the data in the study and accepted responsibility to submit for publication. All authors revised the manuscript and approved the final manuscript as submitted.

Conflicts of Interest

None declared.

References

1. Mensah GA, Fuster V, Murray CJL, Roth GA, Global Burden of Cardiovascular Diseases and Risks Collaborators. Global burden of cardiovascular diseases and risks, 1990-2022. J Am Coll Cardiol 2023 Dec 19;82(25):2350-2473. [doi: 10.1016/j.jacc.2023.11.007] [Medline: 38092509]
2. Crea F. The burden of cardiovascular risk factors: a global perspective. Eur Heart J 2022 Aug 7;43(30):2817-2820. [doi: 10.1093/eurheartj/ehac430]
3. Liu S, Li Y, Zeng X, et al. Burden of cardiovascular diseases in China, 1990-2016: findings from the 2016 Global Burden of Disease Study. JAMA Cardiol 2019 Apr 1;4(4):342-352. [doi: 10.1001/jamacardio.2019.0295] [Medline: 30865215]
4. Roth GA, Mensah GA, Fuster V. The global burden of cardiovascular diseases and risks: a compass for global action. J Am Coll Cardiol 2020 Dec 22;76(25):2980-2981. [doi: 10.1016/j.jacc.2020.11.021] [Medline: 33309174]
5. Goyal P, Kwak MJ, Al Malouf C, et al. Geriatric cardiology: coming of age. JACC Adv 2022 Aug;1(3):100070. [doi: 10.1016/j.jacadv.2022.100070] [Medline: 37705890]
6. Wan X, Ren H, Ma E, Yang G. Mortality trends for ischemic heart disease in China: an analysis of 102 continuous disease surveillance points from 1991 to 2009. BMC Public Health 2017 Jul 25;18(1):52. [doi: 10.1186/s12889-017-4558-3] [Medline: 28743245]
7. The Lancet. Population ageing in China: crisis or opportunity? Lancet 2022 Nov;400(10366):1821. [doi: 10.1016/S0140-6736(22)02410-2]

8. He X, Song M, Qu J, et al. Basic and translational aging research in China: present and future. Protein Cell 2019 Jul;10(7):476-484. [doi: 10.1007/s13238-019-0617-0]
9. Bjarnason-Wehrens B. Elderly patients with ischaemic heart disease need our attention!. Eur J Prev Cardiol 2018 Feb;25(3):276-277. [doi: 10.1177/2047487317750497] [Medline: 29313365]
10. Nanna MG, Chen ST, Nelson AJ, Navar AM, Peterson ED. Representation of older adults in cardiovascular disease trials since the inclusion across the lifespan policy. JAMA Intern Med 2020 Nov 1;180(11):1531-1533. [doi: 10.1001/jamainternmed.2020.2750] [Medline: 32897289]
11. Forman DE, Maurer MS, Boyd C, et al. Multimorbidity in older adults with cardiovascular disease. J Am Coll Cardiol 2018 May 15;71(19):2149-2161. [doi: 10.1016/j.jacc.2018.03.022] [Medline: 29747836]
12. Mendieta G, Pocock S, Mass V, et al. Determinants of progression and regression of subclinical atherosclerosis over 6 years. J Am Coll Cardiol 2023 Nov 28;82(22):2069-2083. [doi: 10.1016/j.jacc.2023.09.814] [Medline: 37993199]
13. Cho SMJ, Koyama S, Honigberg MC, et al. Genetic, sociodemographic, lifestyle, and clinical risk factors of recurrent coronary artery disease events: a population-based cohort study. Eur Heart J 2023 Sep 21;44(36):3456-3465. [doi: 10.1093/eurheartj/ehad380] [Medline: 37350734]
14. Guo Z, Li P, Ge J, Li H. SIRT6 in aging, metabolism, inflammation and cardiovascular diseases. Aging Dis 2022 Dec 1;13(6):1787-1822. [doi: 10.14336/AD.2022.0413] [Medline: 36465178]
15. Yan Z, Wang X, Liu YS, Xing XW, Zhang XG, Lu QH. Decreased serum SIRT6 as a novel predictor of coronary artery disease. Eur Rev Med Pharmacol Sci 2021 Nov;25(21):6660-6669. [doi: 10.26355/eurrev_202111_27110] [Medline: 34787870]
16. Piccininni M, Rohmann JL, Huscher D, et al. Performance of risk prediction scores for cardiovascular mortality in older persons: external validation of the SCORE OP and appraisal. PLoS One 2020;15(4):e0231097. [doi: 10.1371/journal.pone.0231097] [Medline: 32271825]
17. Rao SV, Stone GW. Arterial access and arteriotomy site closure devices. Nat Rev Cardiol 2016 Nov;13(11):641-650. [doi: 10.1038/nrcardio.2016.133] [Medline: 27558003]
18. Rampidis GP, Benetos G, Benz DC, Giannopoulos AA, Buechel RR. A guide for Gensini Score calculation. Atherosclerosis 2019 Aug;287:181-183. [doi: 10.1016/j.atherosclerosis.2019.05.012] [Medline: 31104809]
19. Han JW, Lee SK, Kwon JH, et al. A machine learning algorithm facilitates prognosis prediction and treatment selection for Barcelona clinic liver cancer stage C hepatocellular carcinoma. Clin Cancer Res 2024 Jul 1;30(13):2812-2821. [doi: 10.1158/1078-0432.CCR-23-3978] [Medline: 38639918]
20. Guo J, Wang Z, Wu J, et al. Endothelial SIRT6 is vital to prevent hypertension and associated cardiorenal injury through targeting Nkx3.2-GATA5 signaling. Circ Res 2019 May 10;124(10):1448-1461. [doi: 10.1161/CIRCRESAHA.118.314032] [Medline: 30894089]
21. Zhu Y, Hu S, Pan X, et al. Hepatocyte Sirtuin 6 protects against atherosclerosis and steatohepatitis by regulating lipid homeostasis. Cells 2023 Aug 5;12(15):2009. [doi: 10.3390/cells12152009] [Medline: 37566087]
22. Zhang Q, Tu W, Tian K, et al. Sirtuin 6 inhibits myofibroblast differentiation via inactivating transforming growth factor - β1/Smad2 and nuclear factor - κB signaling pathways in human fetal lung fibroblasts. J of Cell Biochem 2019 Jan;120(1):93-104. [doi: 10.1002/jcb.27128]
23. Pastromas S, Terzi AB, Tousoulis D, Koulouris S. Postprandial lipemia: an under-recognized atherogenic factor in patients with diabetes mellitus. Int J Cardiol 2008 May 7;126(1):3-12. [doi: 10.1016/j.ijcard.2007.04.172] [Medline: 17689745]
24. McAuley PA, Artero EG, Sui X, Lavie CJ, Almeida MJ, Blair SN. Fitness, fatness, and survival in adults with prediabetes. Diabetes Care 2014 Feb;37(2):529-536. [doi: 10.2337/dc13-1347] [Medline: 24062333]
25. Ference BA, Graham I, Tokgozoglu L, Catapano AL. Impact of lipids on cardiovascular health: JACC health promotion series. J Am Coll Cardiol 2018 Sep 4;72(10):1141-1156. [doi: 10.1016/j.jacc.2018.06.046] [Medline: 30165986]
26. Tao R, Xiong X, DePinho RA, Deng CX, Dong XC. FoxO3 transcription factor and Sirt6 deacetylase regulate low density lipoprotein (LDL)-cholesterol homeostasis via control of the proprotein convertase subtilisin/kexin type 9 (Pcsk9) gene expression. J Biol Chem 2013 Oct 11;288(41):29252-29259. [doi: 10.1074/jbc.M113.481473] [Medline: 23974119]
27. Mostoslavsky R, Chua KF, Lombard DB, et al. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell 2006 Jan 27;124(2):315-329. [doi: 10.1016/j.cell.2005.11.044] [Medline: 16439206]
28. Wu X, Liu H, Brooks A, et al. SIRT6 mitigates heart failure with preserved ejection fraction in diabetes. Circ Res 2022 Nov 11;131(11):926-943. [doi: 10.1161/CIRCRESAHA.121.318988] [Medline: 36278398]
29. Borén J, Williams KJ. The central role of arterial retention of cholesterol-rich apolipoprotein-B-containing lipoproteins in the pathogenesis of atherosclerosis: a triumph of simplicity. Curr Opin Lipidol 2016 Oct;27(5):473-483. [doi: 10.1097/MOL.0000000000000330] [Medline: 27472409]
30. Ference BA, Kastelein JJP, Ray KK, et al. Association of triglyceride-lowering LPL variants and LDL-C-lowering LDLR variants with risk of coronary heart disease. JAMA 2019 Jan 29;321(4):364-373. [doi: 10.1001/jama.2018.20045] [Medline: 30694319]
31. Welsh C, Celis-Morales CA, Brown R, et al. Comparison of conventional lipoprotein tests and apolipoproteins in the prediction of cardiovascular disease. Circulation 2019 Aug 13;140(7):542-552. [doi: 10.1161/CIRCULATIONAHA.119.041149]

32. Campeau L, Enjalbert M, Lespérance J, et al. The relation of risk factors to the development of atherosclerosis in saphenous-vein bypass grafts and the progression of disease in the native circulation. A study 10 years after aortocoronary bypass surgery. N Engl J Med 1984 Nov 22;311(21):1329-1332. [doi: 10.1056/NEJM198411223112101] [Medline: 6333635]
33. Quispe R, Martin SS, Michos ED, et al. Remnant cholesterol predicts cardiovascular disease beyond LDL and ApoB: a primary prevention study. Eur Heart J 2021 Nov 7;42(42):4324-4332. [doi: 10.1093/eurheartj/ehab432]
34. Bornfeldt KE, Tabas I. Insulin resistance, hyperglycemia, and atherosclerosis. Cell Metab 2011 Nov 2;14(5):575-585. [doi: 10.1016/j.cmet.2011.07.015] [Medline: 22055501]
35. Lee SB, Ahn CW, Lee BK, et al. Association between triglyceride glucose index and arterial stiffness in Korean adults. Cardiovasc Diabetol 2018 Mar 21;17(1):41. [doi: 10.1186/s12933-018-0692-1] [Medline: 29562908]
36. Irace C, Carallo C, Scavelli FB, et al. Markers of insulin resistance and carotid atherosclerosis. A comparison of the homeostasis model assessment and triglyceride glucose index. Int J Clin Pract 2013 Jul;67(7):665-672. [doi: 10.1111/ijcp.12124] [Medline: 23758445]
37. Vega GL, Barlow CE, Grundy SM, Leonard D, DeFina LF. Triglyceride-to-high-density-lipoprotein-cholesterol ratio is an index of heart disease mortality and of incidence of type 2 diabetes mellitus in men. J Investig Med 2014 Feb;62(2):345-349. [doi: 10.2310/JIM.0000000000000044] [Medline: 24402298]

Abbreviations

AIP: atherogenic index of plasma
ApoA1: apolipoprotein A1
ApoB: apolipoprotein B
ApoE: apolipoprotein E
AUROC: area under the receiver operating characteristic curve
CAD: coronary artery disease
CAG: coronary angiography
CAS: coronary atherosclerosis
DCA: decision curve analysis
FBG: fasting blood glucose
HbA_{1c}: hemoglobin A_{1c}
LAD: left anterior descending branch
LDL-C: low-density lipoprotein cholesterol
LR: logistic regression
***PCSK9*:** proprotein convertase subtilisin/kexin type 9
ROC: receiver operating characteristic
SIRT6: silent information regulator 6
TyG: triglyceride glucose
XGBoost: Extreme Gradient Boosting

Edited by Y Jiang; submitted 17.07.24; peer-reviewed by D Hu, XM Tian; revised version received 27.09.24; accepted 04.10.24; published 15.01.25.

Please cite as:
You Y, Liang W, Zhao Y
Development and Validation of a Predictive Model Based on Serum Silent Information Regulator 6 Levels in Chinese Older Adult Patients: Cross-Sectional Descriptive Study
JMIR Aging 2025;8:e64374
URL: https://aging.jmir.org/2025/1/e64374
doi:10.2196/64374

Original Paper

Wearable Smartphone-Based Multisensory Feedback System for Torso Posture Correction: Iterative Design and Within-Subjects Study

Amanda Polin Pereira[1*], BSc, MSc; Olibario Jose Machado Neto[2*], BSc, MSc, PhD; Valeria Meirelles Carril Elui[1*], BSc, MSc, DSC, HDip; Maria da Graca Campos Pimentel[3*], BSc, MSc, DPhil, HDip, Prof Dr

[1]Faculty of Medicine of Ribeirão Preto, University of São Paulo, Ribeirão Preto SP, Brazil

[2]State University of Mato Grosso do Sul, Nova Andradina MS, Brazil

[3]Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos SP, Brazil

[*]all authors contributed equally

Corresponding Author:
Maria da Graca Campos Pimentel, BSc, MSc, DPhil, HDip, Prof Dr
Institute of Mathematics and Computer Sciences
University of São Paulo
Avenida Trabalhador São-carlense, 400
1203
São Carlos SP, 13566-590
Brazil
Phone: 55 1633739671
Email: mgp@icmc.usp.br

Abstract

Background: The prevalence of stroke is high in both males and females, and it rises with age. Stroke often leads to sensor and motor issues, such as hemiparesis affecting one side of the body. Poststroke patients require torso stabilization exercises, but maintaining proper posture can be challenging due to their condition.

Objective: Our goal was to develop the Postural SmartVest, an affordable wearable technology that leverages a smartphone's built-in accelerometer to monitor sagittal and frontal plane changes while providing visual, tactile, and auditory feedback to guide patients in achieving their best-at-the-time posture during rehabilitation.

Methods: To design the Postural SmartVest, we conducted brainstorming sessions, therapist interviews, gathered requirements, and developed the first prototype. We used this initial prototype in a feasibility study with individuals without hemiparesis (n=40, average age 28.4). They used the prototype during 1-hour seated sessions. Their feedback led to a second prototype, which we used in a pilot study with a poststroke patient. After adjustments and a kinematic assessment using the Vicon Gait Plug-in system, the third version became the Postural SmartVest. We assessed the Postural SmartVest in a within-subject experiment with poststroke patients (n=40, average age 57.1) and therapists (n=20, average age 31.3) during rehabilitation sessions. Participants engaged in daily activities, including walking and upper limb exercises, without and with app feedback.

Results: The Postural SmartVest comprises a modified off-the-shelf athletic lightweight compression tank top with a transparent pocket designed to hold a smartphone running a customizable Android app securely. This app continuously monitors sagittal and frontal plane changes using the built-in accelerometer sensor, providing multisensory feedback through audio, vibration, and color changes. Patients reported high ratings for weight, comfort, dimensions, effectiveness, ease of use, stability, durability, and ease of adjustment. Therapists noted a positive impact on rehabilitation sessions and expressed their willingness to recommend it. A 2-tailed t-test showed a significant difference (P<.001) between the number of the best-at-the-time posture positions patients could maintain in 2 stages, without feedback (mean 13.1, SD 7.12) and with feedback (mean 4.2, SD 3.97), demonstrating the effectiveness of the solution in improving posture awareness.

Conclusions: The Postural SmartVest aids therapists during poststroke rehabilitation sessions and assists patients in improving their posture during these sessions.

(JMIR Aging 2025;8:e55455) doi:10.2196/55455

KEYWORDS

stroke rehabilitation; posture; postural balance; wearable technology; multisensory feedback; smartphone; stroke; mHealth; mobile health; digital health; digital technology; digital intervention; wearable technology; gerontology

Introduction

The 2019 Global Burden of Disease report indicates that stroke is a major global health issue, with around 12.2 million new cases and 6.6 million deaths worldwide that year. This makes stroke the second leading cause of death and the third leading cause of disability [1], with its occurrence increasing with age and affecting both males and females [2]. In Brazil, stroke is the leading cause of death, and mortality from stroke has increased in recent years. Registering 103,000 deaths in 2019, the number rose to over 112,000 by 2023 [3].

One of the most common and challenging poststroke conditions is hemiparesis, which affects a substantial number of stroke survivors globally [4,5]. For these patients, controlling trunk movement becomes a fundamental motor skill essential for performing various functional tasks [6,7]. For instance, research shows that maintaining proper posture correlates with walking ability in patients undergoing acute stroke rehabilitation [8].

In individuals without hemiparesis, posture typically involves symmetric and balanced alignment of all body parts, including the head, shoulders, spine, hips, and limbs, following the body's natural curves. In contrast, a person with hemiparesis, characterized by muscle weakness on one side of the body due to brain injuries or stroke, may exhibit an asymmetric posture, resulting in an imbalanced or tilted position. We refer to the term "best-at-the-time posture" to describe the optimal posture achievable by a hemiparetic patient, considering their motor limitations (we use the term "correct posture" interchangeably). The goal is to achieve the highest level of alignment and balance possible.

The literature registers many efforts to support stroke survivors with respect to improving trunk stability [9], trunk compensation [10], motor control [11-13], and accessing mobility [14]. Furthermore, a systematic review has confirmed the significantly positive impact of trunk training on various aspects of trunk control, sitting and standing balance, and mobility [15].

In health care, literature acknowledges the promise of smartphones [14,16] and wearables [17-20] in empowering individuals, aiding diagnosis, promoting behavior change, and enabling self-monitoring. Wearable and rehabilitation devices for different body parts, such as head, limbs, and torso, enhance training outcomes through valuable feedback [21-24]. Moreover, numerous studies have affirmed the feasibility of inertial sensors for balance and gait assessment [25-28], as corroborated by comprehensive literature reviews [29,30], even for individuals with chronic stroke [31].

In the context of wearable technology for poststroke patient support, researchers have delved into various aspects. For instance, studies demonstrated the feasibility of using step activity monitors for patients with recent stroke [32,33]. One study involved the use of intelligent insoles for analyzing the gait of patients with hemiparesis [34]. Another result is a home-based rehabilitation system using wearable sensors in smartwatches, allowing therapists to monitor patients remotely [35]. In another study, the authors evaluated the impact of haptic nudging delivered via a wrist-worn wearable device on upper limb movement during inpatient stroke rehabilitation [36]. Additionally, a recent study reported a test for assessing kinematic parameters in chronic stroke survivors. This test used a standardized mobility assessment with a simple smartphone attached to the lumbar spine using an elastic band to measure participants' kinematics [14].

Existing literature explores a range of wearable and mobile device-based solutions for postural monitoring. In the context of hemiparesis, researchers have investigated the effects of rhythmic haptic cueing on spatial and temporal gait characteristics using haptic feedback [37]. In a broader context, Smart Pose leverages a smartphone's camera, accelerometer, and magnetometer to detect poor neck posture during smartphone use, providing feedback through vibrations, text messages, and alarms [38]. Additionally, a 3-axis accelerometer biofeedback system corrected neck posture during prolonged computer use and effectively reduced inappropriate neck angles [28]. Other innovations include elastic t-shirts with embedded sensors for posture feedback [39], wearable devices for spine posture monitoring [40], and posture differentiation systems [41].

However, it is essential to recognize that effective functional rehabilitation requires both body awareness and torso control to ensure upper limb functionality [4,7,9,42]. Additionally, concerns regarding the cost and availability of treatment [43-46] highlight the need to explore alternatives that use ubiquitous, low-cost smartphones. Despite the growing number of mHealth mobile apps for patients with stroke [47], few address trunk control as a primary focus [27,48,49]. Furthermore, while both intrinsic and extrinsic feedback are crucial for motor learning after stroke [50-52], and therapists spend a significant amount of time providing this feedback [53,54], few efforts used multisensory feedback from smartphones for upper body rehabilitation [24]. Moreover, while smartphone-based multisensory feedback has proven effective for postural monitoring in healthy individuals [55], similar solutions for patients with stroke are still lacking. This highlights the need for research into new wearable technologies to enhance rehabilitation support and improve posture and trunk awareness in patients with stroke.

To address this challenge, we aimed to develop Postural SmartVest. This affordable wearable technology takes advantage of low-cost smartphone resources by (1) leveraging the built-in accelerometer sensor to monitor sagittal and frontal plane changes continuously and (2) exploiting the device's visual, tactile, and auditory feedback to guide patients in performing the movements required to return to their best posture at the time.

Methods

Overview

The research group comprises 2 occupational therapists (authors APP and VMCE) and 2 computer scientists (authors OJMN and MDGCP). To design the initial software architecture, the group received input from other health and computing professionals [56].

The study was conducted in a specialized rehabilitation center that requires both ethical committee approval and verification of ethical and safety aspects for studies conducted on its premises. Therapists participating had at least 1 year of experience in stroke rehabilitation and received training on device use before working with patients. The smartphones used in the feasibility study were either from the research team, having been tested and shown to function properly, or were participants' own devices. To minimize malfunction risks, no smartphone was allowed to be charged during use by participants. The Informed Consent Form informed participants about potential discomfort due to heat and the vest and assured them that, in the unlikely event of a smartphone malfunction causing an explosion, immediate assistance and emergency services would be provided. Data acquisition and storage were handled securely, with all data kept anonymized and stored in accounts requiring login via a secure network. Photographs were authorized with face identification removed. Participants consented to the use of data for the study and were informed of their right to withdraw at any time without coercion or obligation. Participation was voluntary and did not include any financial benefit, in compliance with Brazilian regulations.

The study comprised design and development iterations involving brainstorming meetings, interviews with therapists and patients, prototype development, feasibility studies, and experimental sessions (see workflow in Textbox 1).

Textbox 1. Research workflow.

Literature review

Iteration 1

- Brainstorm meetings
- Interview with therapists
- Prototype #1
- Feasibility study

Iteration 2

- Prototype #2
- Pilot study

Iteration 3

- Prototype #3
- Kinematic assessment
- Postural SmartVest

Within-subjects study

- Presession interviews with patients
- Experimental session
- Postsession interviews with patients
- Postsession interviews with therapists

Throughout the study, our literature review was ongoing and dynamically aligned with emerging themes. Initially, it concentrated on the specific needs of patients with stroke [4] and the effectiveness of trunk training [15]. Next, we reviewed work on novel wearable [21] and affordable smartphone-based solutions [31], including reports in literature reviews [29,57], and noticed the involvement of healthy adults in feasibility studies [26,58], as well as control participants [27]. These works are representative of the broader ongoing review process, which continued to cover these and additional relevant areas throughout the study.

We used the Google Sheets software for statistical analysis relative to the within-subject study. We used thematic analysis [59] for categorizing interview responses. For the kinematic assessment, we compared angular momentum from values calculated by Vicon and the smartphone’s sensors.

Participants

Our study involved 3 groups: poststroke participants, therapists, and healthy participants. We used interviews and questionnaires to collect requirements, demographic data, and their impressions of Postural SmartVest.

For the interviews with therapists, the inclusion criterion was having at least 1 year of experience in stroke rehabilitation. We invited professionals from a rehabilitation center using printed flyers and email invitations. In the feasibility study, we recruited participants via university email: the inclusion criteria were being 18 years or older and owning a smartphone. The exclusion criterion was self-reported physical or cognitive impairments. For the pilot and within-subjects studies, the inclusion criteria were having chronic hemiparesis, being 18 years or older, walking independently, having cognitive ability for communication, and attending rehabilitation sessions at least twice a week for a minimum of 1 month.

Iteration 1: Brainstorming Meetings, Interviews With Therapists, and Feasibility Study

For the brainstorming meetings, the research team held three 2-hour sessions to gather functional and nonfunctional requirements for the solution to develop the initial prototype comprising the vest and the app, Prototype #1 (Figure 1).

For the interviews with therapists, one team member (APP) conducted semi-structured interviews with therapists. The interview adhered to a structured protocol comprising 8 questions categorized into 4 categories: resources for posture improvement, general technology acceptance, specific technology requirements, and smartphone use (Multimedia Appendix 1: App A).

Figure 1. Prototype #1 in use in the deployment phase of the feasibility study.

For the feasibility study, we used the first version of the app. We created 8 vests and conducted sessions with 8 participants at a time. We provided smartphones to participants who could not use their own. Each session lasted 2 hours and followed a 3-phase protocol: preparation, deployment, and follow-up. In the preparation phase, therapists adjusted the vests and calibrated the seating posture for each participant.

In the deployment phase, participants set their devices to airplane mode and wore headphones to ensure privacy for their feedback. They then engaged in smartphone-free activities like reading or using their laptops (Figure 1). The deployment phase consisted of two 30-minute segments: feedback from the app was turned off during the first part and turned on during the second. The app continuously recorded the participants' postural data throughout both segments in this phase.

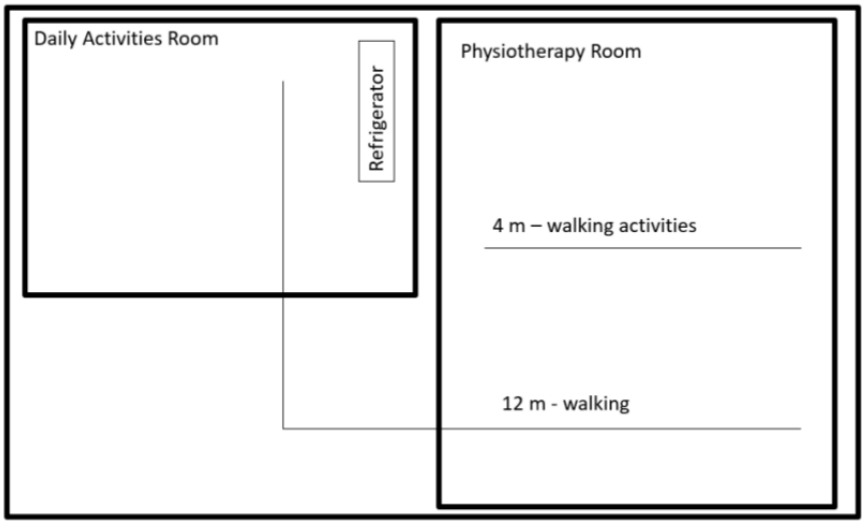
In the follow-up phase, participants responded to 2 questionnaires, one with scaled questions and the other with open-ended inquiries. These questionnaires aimed to assess usability and satisfaction. One questionnaire evaluated Postural SmartVest as an assistive technology for posture correction (Multimedia Appendix 1: App B), while the second assessed Postural SmartVest as an assistive technology in general (Multimedia Appendix 1: App C). For the latter questionnaire, we adapted the Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST; version 2.0), a 12-item assessment designed to measure user satisfaction with devices and services [60]. We used a translated and validated version [61] and focused on the device-related section of the questionnaire. This section included 8 scale questions that assessed size, weight, adjustability, safety, durability, ease of use, comfort, and overall effectiveness of the technology.

Iteration 2: Pilot Study

The feedback from the previous iteration guided the design of Prototype #2. We then used this prototype in a pilot study involving 1 patient with hemiparesis to observe the solution and detail the within-subjects study design. In the pilot study and the following studies, the smartphone used was a Motorola G with Android (version 5.1).

During the pilot study, the poststroke patient followed a protocol that included 5 ambulant-based daily activities conducted in the therapy room (Figure 2). The protocol consisted of 4 distinct phases. In the first phase, the participant walked 8 m, incorporating forward, sideways, and backward walking. In the second phase, they proceeded with a 12-m free walk to a refrigerator. In the third phase, the participant engaged in an upper limb activity that entailed opening the refrigerator, taking a glass, and returning it to the fridge. Finally, in the fourth phase, they completed a 12-m free walk back to the starting point.

Figure 2. Layout of the therapy room.



Iteration 3: Kinematic Assessment

The feedback from the pilot study informed the development of a new prototype (Prototype #3). At this stage, we checked the accelerometer readings obtained by the app running on Motorola G with Android (version 5.1).

On the one hand, sensor type and placement, activity characteristics, and population-specific conditions can influence the accuracy of sensor outcome estimates, as reported in a study that used research-grade wearable sensors on the upper arms, waist, and ankles and included poststroke patients [62]. On the other hand, the literature reports positive results regarding using state-of-the-art smartphone accelerometer technology. For instance, researchers investigated the comparative performance of 3 different commercially available smartphone accelerometers among themselves and to a gold-standard Vicon MX motion capture system, indicating that the devices are valid and reliable measuring instruments for estimating linear accelerations [63].

We assessed the angle calculations of the third prototype by comparing them with the Vicon Gait Plug-in system, a gold-standard in full-body kinematic and kinetic modeling. The Vicon system includes software and fixed cameras that capture signals from a moving target, providing visual feedback through animated vectors and plans, and it also calculates quantitative data, including distances and angles between selected planes.

Under the guidance of one occupational therapist (APP), author OJMN used the prototype to perform 6 movements: trunk flexion, trunk extension, left and right-side bending, and left and right trunk rotation (Figure 3).

The therapist calibrated the subject's optimal posture position, and for all movements, the subject began in an upright position, executed the movement, and returned to the initial position. The Vicon system and the app prototype collected data simultaneously, following an initial manual synchronization process, with the Vicon recording frames at 250 frames per second and the app prototype recording angular moments at each second along with coordinate values. To reconstruct the trunk segment, we placed markers on the clavicle, sternum, cervical vertebra C7, thoracic vertebra T10, and a point located on the medial border of the right scapula. We conducted 6 sessions using 39 markers on a full-body model (Figure 4).

Following the kinematic assessment, Prototype #3 was referred to as SmartVest and was used for the remainder of the study (Figure 5).

Figure 3. Movements used for simultaneous data collection by Vicon and the app.

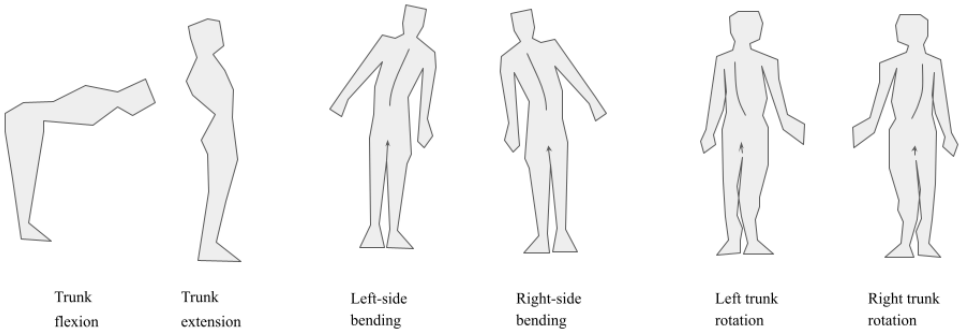


Figure 4. Simultaneous data collection by Vicon and the app.

Figure 5. The Postural SmartVest solution comprises a customizable smartphone app placed in a transparent pocket on an athletic, lightweight compression tank top.

Within-Subjects Study

The within-subjects study included individual sessions with poststroke patients, each accompanied by their therapists. Before that, therapists used the Postural SmartVest for about 30 minutes while the researchers provided guidelines to ensure the professionals understood how the solution worked. Only afterward did the therapists apply Postural SmartVest to their patients during their rehabilitation sessions. All sessions used the same smartphone (Motorola G with Android; version 5.1).

In the presession interview with each patient, we gathered information about their posture control during activities, awareness of posture issues, and any potential biases toward one side of the body. We also inquired about their strategies to correct posture problems when they noticed them (Multimedia Appendix 1: App D).

At the beginning of each session, the therapist adjusted the vest on the patient, calibrated the app to recognize the patient's best-at-the-time standing posture, and initiated the monitoring process (Textbox 2). Subsequently, the patient performed the activities defined in our pilot study twice, in consecutive sessions, each with an estimated duration of 10 minutes (Figures 6 and 7). The app did not provide feedback in the first session but registered all the movements. In the subsequent session, the app provided feedback through screen color changes (green and red), vibrations, and audio guidance.

Afterward, we asked the poststroke patients to fill out 2 questionnaires, one focusing on Postural SmartVest as an assistive technology for posture correction (Multimedia Appendix 1: App B) and the other assessing Postural SmartVest as a general assistive technology (Multimedia Appendix 1: App C).

Additionally, the therapists shared their feedback after working with the patients using a tailored version of the Postural SmartVest QUEST (version 2.0) questionnaire (Multimedia Appendix 1: App E). This study helped us understand the effectiveness of Postural SmartVest during stroke rehabilitation, both from the patient's and therapist's perspectives.

Textbox 2. Calibration workflow.

Adjust the vest on patient

Start the app

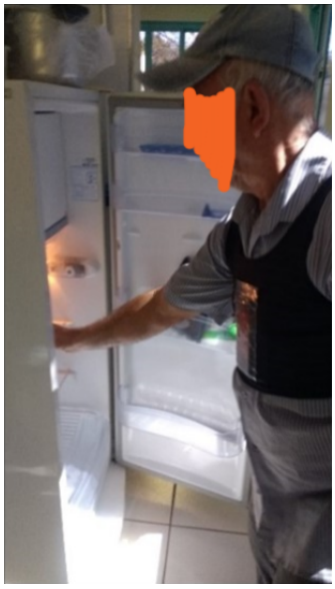
Access settings

- Enter the tolerance threshold angle for frontal movement (in degrees)
- Enter the tolerance threshold angle for lateral movement (in degrees)
- Enter the allowable duration for temporary deviations from the calibrated posture (in seconds)
- Ensure the option for vibration feedback is selected
- Ensure the option for audio feedback is selected
- Guide the patient to assume his best posture at that moment
- Press the "Calibrate" button to configure the current position as the best posture for monitoring
- Return to main screen

Press play or pause to initiate monitoring

Figure 6. Walking activity.

Figure 7. Upper limb activity.



Ethics Approval

The study received prior approval from the Hospital of the Clinics of the Ribeirão Preto School of Medicine at the University of São Paulo, Brazil, identifier 57234816.3.0000.5440. As detailed in the text approved by the ethics committee, the study adhered to ethical principles by ensuring beneficence through rigorous monitoring and support for participants, justice by offering equitable access and fair treatment, and respect by obtaining informed consent and addressing potential discomfort and safety concerns proactively.

Results

Participants

For the interviews, we enlisted 20 therapists: 10 occupational therapists and 10 physiotherapists. All participants were female, aged 26-40, with varying years since graduation (Table 1).

For the feasibility study, we recruited 28 healthy participants, aged 20-62, with an average age of 30.9. The group included 17 males and 11 females (Table 1). Of the participants, 25 used their own smartphones, while 3 used devices provided by us. The study covered 22 different smartphone models and 7 versions of the Android operating system.

For the pilot and within-subjects studies, we recruited poststroke participants from patients attending rehabilitation sessions at least twice a week for a month at the exact center. Inclusion criteria included chronic hemiparesis from stroke, aged 18 years or older, independent walking ability, cognitive ability for communication, and participation in rehabilitation sessions for at least a month.

We enrolled 1 patient for the pilot study and 40 patients for the within-subjects study (27 males, 67.5%), with a mean age of 57.1 (SD 11.42) years, and functional independence categorized as independent (n=39) of moderate assistance (n=1) on the functional independence measure [64,65]. Among them, 24 had right hemispheric lesions, 16 had left hemispheric lesions, and 38 had ischemic strokes. The duration of their condition ranged from 6 months to 19 years, with a majority between 1 and 4 years (Table 1).

Table 1. Participants' demographics.

Characteristics	Values
Therapists (n=20; all female), n	
Age range (years)	
26-30	8
31-40	12
Years since graduation[a]	
1-2	2
3-5	6
6-9	9
10+	3
Healthy participants (n=28; 17 male)	
Age (years)	
Range	20-62
Mean	30.9
Smartphone used, n	
Participant’s own	25
Ours	3
Poststroke patients in the within-subjects study (n=40; 27 male)	
Age range (years)	
20-50	8
51-60	16
61-70	11
71-80	5
Time since stroke (months; mean 43.9 months), n	
0-12	16
13-24	5
25-36	4
37-48	5
48+	10
Affected hemisphere	
Right	24
Left	16
Type of stroke	
Ischemic	38
Hemorrhagic	2
FIM[b]	
Moderate assistance	1
Moderate or complete independence	39

[a]Having at least 1 year of experience in stroke rehabilitation.

[b]Functional independence measure.

Iterative Design

The Postural SmartVest solution comprises a customizable smartphone app for Android devices securely placed in a transparent pocket on an athletic, lightweight compression tank top (Figure 5).

Vest Component Evolution

The development of the vest component went through several iterations (Figure 8), with 6 prototypes created based on therapist's input (see responses in Multimedia Appendix 1: App F-G) and feedback gathered during the feasibility study (see responses in the Multimedia Appendix 1: App H). The final version of the garment used in the Postural SmartVest uses an off-the-shelf lightweight compression tank top made from polyamide with elastane. We added a perforated plastic pocket to the top front. The pocket serves 3 primary purposes: securely holding the smartphone, allowing for easy headset connection, and facilitating heat dissipation.

Figure 8. Vest prototypes developed during the iterative design.

App Component Evolution

The smartphone app also underwent significant changes during the iterative design process. Initially conceptualized following brainstorming sessions that defined both functional and nonfunctional requirements (Textbox 3 and Multimedia Appendix 1: App F), the app's preliminary software architecture was informed by input from health and computing professionals [56].

Overall, the app was guided by simplicity, feasibility, and adaptability principles, in line with cost-effective assistive technology solutions from our previous research [66]. Refinements and improvements were made based on therapist interviews and feedback obtained during the feasibility study

(Textbox 4). The final version has a main screen and a sliding panel for configuration options (Figure 9).

The app records every activity along with its corresponding timestamp. This record encompasses posture changes, feedback delivery, and adjustments to application settings. During the design, we evolved the app to use schemes suggested by the OpenMHealth organization [67] when saving log data.

Textbox 3. App requirements summary.

Functional requirements

- Therapist can calibrate best-at-the-time posture with 1 touch.
- Therapist defines threshold values for small movements.
- Therapist defines threshold values to indicate how long the patient can be off best-at-the-time posture before the app starts guiding the patient.
- The app guides the patient back to their calibrated best-at-the-time posture using screen color, vibration, and audio messages.
- The app must record and store all interactions for later export and analysis.

Nonfunctional requirements

- The app should work on a low-end smartphone.
- The smartphone should be attached to a vest within a transparent pocket.
- The app should be compatible with Android, the most common operating system in Brazil.
- The app should work with the screen off to conserve battery if needed.

Textbox 4. App development iterations

Prototype #1:

- Implements all functional and nonfunctional requirements
- Features a single screen displaying x, y, and z coordinates, directions, settings, and a calibration button
- Calibration settings are not stored for future use
- Provides 5 guiding messages in Brazilian Portuguese (1 for each direction and 1 for achieving the best-at-the-time posture)

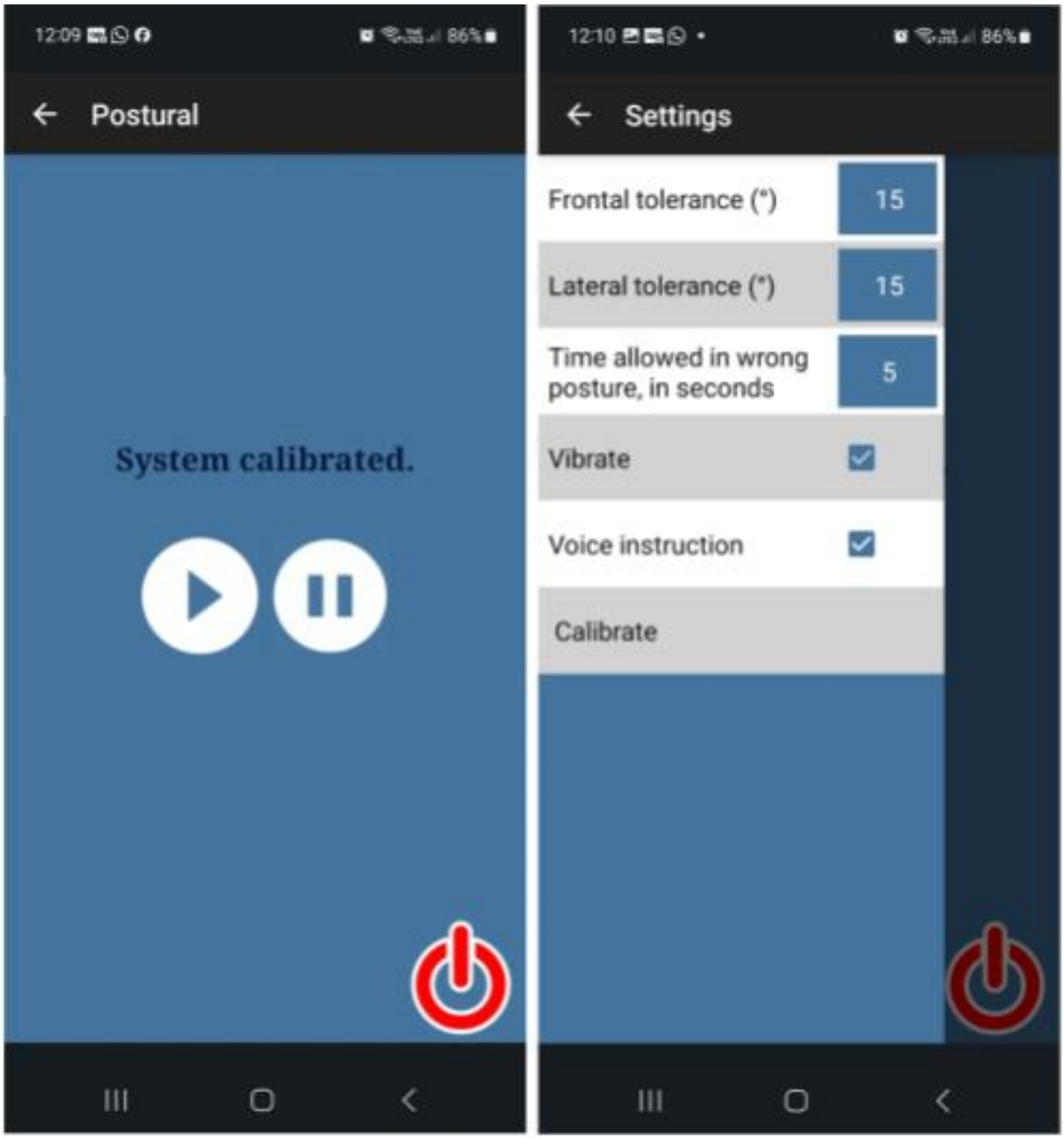
Prototype #2:

- Replaces coordinates and directions with information on sagittal and frontal planes, using values in degrees
- Introduces a new configuration screen for settings and calibration
- Stores and uses calibration settings until a new calibration is performed
- Stores all coordinate readings and feedback provided (version for kinematic assessment)
- Includes play and pause buttons for monitoring, along with an exit button on the home screen
- Expands the number of audio instructions, available in both English and Spanish
- Offers audio feedback every 5 minutes of maintaining the best-at-the-time posture

Prototype #3:

- Information on planes renamed to frontal and lateral planes
- Includes a help option with instructions on how to use the app

Figure 9. App main screen (left) and configuration panel (right).



App Component Interface

The app's main screen provides large buttons for easy access to pausing, resuming, and closing the application (Figure 9, left side). By clicking on "Postural" in the top bar on the left side of Figure 9, therapists can access 6 customization settings, as shown on the figure's right side. From top to bottom in the figure, the first 5 customization parameters allow therapists to configure thresholds for frontal and lateral movements (in degrees), to define the allowable duration for temporary deviations from the calibrated posture (in seconds), and to personalize feedback to the patient, including options for vibration and voice instructions. The therapist uses the last option, Calibrate, to set the best-at-the-time posture that the patient with hemiparesis can achieve at that moment. When calibration is activated, the app registers the current coordinates as the target posture for the patient to achieve and provides audio feedback.

After calibrating, the therapist can return to the main screen using the arrow in the top bar. From that point onward, the therapist can use the play and pause buttons to activate or deactivate posture monitoring. Additionally, we have included a large "off" button to allow the therapist to exit the application easily.

Monitoring starts when the therapist hits the play button, and the whole screen changes to green. If a patient's posture deviation exceeded the predefined thresholds for lateral and frontal angles relative to the calibrated posture, the app responded immediately with visual feedback, transitioning the screen color from green to red, and tactile feedback through vibrations. Furthermore, the app introduced audio feedback if a patient deviated from the best-at-the-time posture beyond the

allowed time for temporary deviations from the calibrated posture. Additionally, the app gives positive feedback every 5 minutes if the patient is maintaining the best-at-the-time posture. This comprehensive feedback system provided patients with real-time guidance tailored to their unique needs and specific postural challenges.

The app provides audio instructions: “please lean forward,” “please lean backward,” “please lean to the right,” and “please lean to the left.” We implemented enhancements throughout the iterations, such as adding positive feedback messages to encourage users to maintain the best-at-the-time posture and random variations in the messages that congratulate users for returning to their best-at-the-time posture.

To meet a requirement identified during the feasibility study, where participants used devices set to various languages, we improved the app to support users in English, Spanish, and Portuguese. It dynamically adapts its interface and voice instructions based on the language settings of the user’s device.

Kinematic Assessment

The kinematic assessment results, which involved comparing angular momentum values calculated by the Vicon system with those obtained from the smartphone's sensors, are presented graphically in the Multimedia Appendix 1: App I. The results consistently demonstrated that smartphone accelerometers offer reliable data readings, and our mobile application accurately interprets these values. These outcomes align with existing literature [24,28,38,40,58], reinforcing the utility of our approach within the context of our application. Moreover, this positive evaluation of the system's trustworthiness is in line with the feedback received from therapists regarding its performance and reliability, as detailed in the section *Postsession Interviews With Therapists*.

Pilot Study

The pilot study confirmed the proposed activity protocol for the within-subjects study, with participants safely completing the walking circuit within the estimated time frame.

Within-Subjects Study

Presession Interviews With Patients

Findings are summarized in Table 2. Out of the 40 participants with hemiparesis, 34 (85%) acknowledged experiencing challenges in balancing one side of their body. These participants reported various triggers for posture correction, including pain, the sensation of weight, and fatigue. Specifically, 6 participants stated that they only adjusted their posture when they felt pain, 3 did so in response to a sense of heaviness on the side affected by hemiparesis, and 2 did it to alleviate fatigue. Participants mentioned that they recognized the need for correction only when viewing photos (2 participants), standing in front of a mirror (2 participants), or following verbal advice from a family member (6 participants). In total, 4 participants admitted that they often neglected to correct their posture due to forgetfulness, with one of them explaining, “It is hard to remember, and correcting myself all day is tiring.”

When providing opinions on assistive technologies, participants considered the technologies “viable for use” and noted that they provide “progress to believe in improvement.” They also mentioned that these technologies “make rehabilitation more precise” and “help more than just me and the therapist.” Additionally, it was observed that they “stimulate people's lives even with minimal gain.” The repetition of these responses highlights the consistent perception of the technologies' value and potential impact.

Regarding the necessary requirements for the technology, 11 participants agreed that it could assist in the perception and control of the trunk. They envisioned the technology as a “chair that helps maintain correct posture” or “something that allows for posture visualization.” Some participants suggested a device that “corrects the arm and leg to prevent them from falling” and another that “reminds me to correct my posture.” They also proposed features such as “notifying to stay straight” or having the ability to “pull to the side.” Other suggestions included a “vest to correct posture” or a device that “attaches to the trunk and holds it in place.” Additional requirements included providing a “signal on the body,” “holding the shoulders, reminding to keep the back straight, and being fixed on the bra,” and being “close to the body,” “comfortable, practical, and unobtrusive,” and “adaptable to the body.” Participants emphasized that the technology should be “integrated with therapy,” “usable over time,” and “inexpensive.”

Table 2. Patients' report on posture, assistive technology, and smartphone usage.

Category	Patients, n
Maintenance of correct posture	
Unable to maintain posture	22
Able to maintain posture	17
Able to maintain posture sitting, but not standing	1
Perception of correct posture	
Perceive correct posture	36
Do not perceive correct posture	4
Difficulties and postural correction	
Remembering to correct and trying to adjust	8
Unable to correct posture	6
Need to leave sitting posture to correct	1
Triggers for posture correction	
After perceiving pain	6
After request from family	6
After perceiving weight on the side of hemiparesis	2
After feeling tired	2
After observing photos	2
After seeing oneself in the mirror	2
Opinion on assistive technology	
Believe it can help but consider it expensive	13
Knowledge from TV but no contact	2
Should be offered by the public health system	1
Would be good if recommended by professionals	1
Smartphone usage	
Participants with smartphones	26
Made accessibility adjustments	6
Experience difficulties	9
Use daily	19
Use sporadically	2
Use only for calls	4
Use only for WhatsApp, music, and camera	1
Other uses (WhatsApp, Facebook, YouTube, and internet banking)	12

Session Data

The average duration of the stages with and without feedback was 9.14 (SD 5.36) minutes and 8.52 (SD 3.21) minutes, respectively. A 2-tailed *t*-test indicated that these differences were not statistically significant.

We compared the number of the best-at-the-time posture positions patients could maintain in the 2 stages: without feedback and with feedback. The 2-tailed *t*-test results indicated a statistically significant difference between these 2 conditions. Specifically, in the stage with feedback, patients exhibited a higher ability to maintain correct positions (mean 13.1, SD 7.12) compared to the stage without feedback (mean 4.2, SD 3.97). This significant difference ($P<.001$) underscores that patients could achieve and sustain their best-at-the-time postures more frequently when Postural SmartVest provided feedback.

Additionally, we examined the average number of movements performed by patients to attain their best-at-the-time posture in both the stages without feedback and with feedback. In the stage without feedback, patients averaged 8.09 (SD 6.88) movements. In contrast, in the stage with feedback, they averaged 4.49 (SD 1.49) movements. The 2-tailed *t*-test results indicated no statistically significant difference between the values obtained in these 2 stages, with a *P* value of .11.

Postsession Interviews With Poststroke Patients

In the evaluation of satisfaction with Postural SmartVest, 37 (92.5%) patients reported being satisfied or very satisfied with the support provided by Postural SmartVest to help them maintain proper posture. In total, 2 respondents had neutral feelings, and 1 expressed dissatisfaction with the solution. In total, 39 (97.5%) patients said they would recommend the solution to others with hemiparesis due to stroke, and one indicated that they might recommend it.

When asked if they were able to maintain their correct posture, 17 patients answered “yes,” 2 responded with “sometimes,” and 21 answered “no.” When asked if they noticed that their body often tilted to a particular side, 18 reported that their body tilted to the right side, 16 indicated that their body tilted to the left, 3 stated that their body tilted backward, and 3 could not respond.

Responses to the questionnaire evaluating Postural SmartVest as an assistive technology using a 1-5 scale (Multimedia Appendix 1: App E) are summarized in Table 3. Regarding open-ended questions directed to the patients, we highlight in Table 4 some critical opinions and the number of users who expressed them.

Poststroke patients provided rich feedback on Postural SmartVest’s usability and effectiveness (Table 4). They suggested design improvements, such as incorporating openings with zippers and using breathable materials for cell phone accommodation. Some recommended extended usage periods, while others emphasized early implementation for gradual posture improvement. Software suggestions included customizable message frequency and varied vibration patterns. Participants highlighted the positive impact of audio feedback, likening it to therapist interactions during sessions. These insights, shared by a significant portion of the 40 patients, showcase the multifaceted utility of Postural SmartVest and its potential to address various poststroke rehabilitation needs. The integration of smartphones into thoracic clothing proved safe during our study, with no observed risks during continuous 1-hour use.

Table 3. Patients’ evaluation of postural SmartVest as an assistive technology.

Aspect	Favorable responses, %
Weight	100
Comfort	100
Dimension	97.5
Effectiveness	97.5
Ease of Use	87.5
Stability	82.5
Durability	72.5
Ease of Adjustment	67.5

Table 4. Patients’ responses to open-ended questions. The last column indicates the number of participants reporting.

	Responses, n
“I would suggest improvements to the garment, such as an opening Velcro or zipper and the use of breathable material to accommodate the cell phone.”	10
“The vest should be more discreet so that it could be worn more imperceptibly under the shirt.”	10
“The garment should be worn for a longer period, over one hour.”	4
“The solution should be introduced to me as early as the first rehabilitation session, immediately after a stroke, so I could get used to perceiving and correcting my posture earlier.”	4
“I would like to adjust the audio volume of the messages through the app.”	3
“The frequency of the messages should be customizable.”	3
“I would like different vibration patterns for different body movements.”	3
“Listening to the app's directions reminds me of you (the therapist) talking to me during the rehabilitation sessions.”	1
“I would use the solution to prevent falls.”	1
“The solution helps me walk.”	1
“The solution helps me improve my motor perception.”	1
“I believe that the vest alone can play an important role in postural perception.”	1
“I think that using the solution after some rehabilitation sessions is better because right after the stroke, I could not minimally perceive or correct my posture.”	1

Postsession Interviews With Therapists

Responses to the adapted QUEST (version 2.0) questionnaire evaluating the solution as an assistive technology (Multimedia Appendix 1: App E) are summarized in Table 5. In total, 2 therapists provided a neutral rating for "Support to Posture" and "Audio guidance accuracy." They explained that patients with spatial orientation difficulties still required assistance from therapists despite using the solution. In total, 5 therapists provided neutral responses regarding "Solution trustworthiness." In total, 2 therapists recommended that the app recognize other body parts and suggested improvements in the instructions, noting that the app reinforced the user's reference sides.

We summarize the responses to evaluating satisfaction with Postural SmartVest on a 1-10 scale (Multimedia Appendix 1: App E) in Table 6. The responses indicated a highly positive impact on patients' posture control, with 95% rating it as "A lot." Additionally, 90% of the therapists believed that the app significantly contributed to achieving therapy goals. Moreover, 60% reported conducting fewer patient posture-related interventions after using the app. In terms of ease of use, 70% of the therapists found it "Very easy" for patients to follow the app's guidance.

Table 5. Therapists' responses to the adapted QUEST (version 2.0) questionnaire (1-5 scale).

Response	Rating 4 or higher, %	Indication
Support to Posture	90	Very satisfied
Audio guidance accuracy	90	Very satisfied
Device comfort	95	Very comfortable
Solution trustworthiness	83	Very reliable
Safety during use	100	Very safe
Likelihood to recommend	100	Very likely

Table 6. Therapists' satisfaction with SmartVest (1-10 scale).

Question	Percentage: answer	Indicating
Did you notice a positive impact on patients' posture control?	95%: 8 or higher	A lot
Do you think the app contributed to the therapy goals?	90%: 8 or higher	A lot
Did you notice if you conducted more or less posture-related interventions for patients?	60%: 3 or less	Fewer
How difficult was it for the patients to follow the app guidance?	70%: 3 or less	Very easy

Discussion

Principal Results

We designed Postural SmartVest, a solution consisting of a chest garment holding a smartphone running an Android application. When the user wears the Postural SmartVest in a rehabilitation session, the therapist first calibrates the app to the best posture the patient can achieve at the time. The app uses the smartphone accelerometer to identify variations in the user's body movement in the sagittal and frontal planes relative to the calibrated position. Based on these variations, the app provides multisensorial feedback that guides the patient to perform the movements required to return to the calibrated position.

The iterative development of Postural SmartVest, informed by therapist interviews and feasibility study feedback, tailored the final product to user needs. The app reliably identifies posture changes and provides corresponding multisensorial feedback, as substantiated by kinematic assessment.

Postural SmartVest significantly improved patients' ability to maintain their best-at-the-time postures, with patients showing a higher capacity for correct positions when using feedback. Although the analysis found no significant difference in the number of movements between stages, the increased movement during the feedback stage suggests an active engagement with the system for postural adjustments. This engagement aligns with rehabilitation goals, emphasizing motor skill enhancement and proprioception. These findings highlight Postural SmartVest's potential as an assistive technology for poststroke patients, emphasizing the importance of feedback in rehabilitation tools.

Postural SmartVest's flexibility meets each patient's unique needs and progress, with a highly customizable app enabling therapists to personalize feedback and adjust thresholds. This multisensory approach enhances engagement and may result in more effective posture correction. The app offers a comprehensive feedback system, including visual cues, tactile vibrations, and audio instructions.

The technology received positive feedback regarding its design and usability, with patients and therapists giving favorable ratings for various aspects of Postural SmartVest, including weight, comfort, dimensions, and effectiveness. Patients believed in its utility for walking assistance and motor perception improvement, and their suggestions aligned with their approval. Many patients were willing to recommend Postural SmartVest to other patients with poststroke hemiparesis, highlighting its perceived value. Healthy users also expressed satisfaction with Postural SmartVest's posture support.

Limitations

Our study used a specific sample of poststroke patients, which may not fully represent the diverse stroke survivor population. This highlights the need for including a broader and more diverse sample in future research. Although the findings indicate high satisfaction among both patients and therapists, individual preferences and responses varied. Patient feedback identified several limitations that suggest potential improvements to the app's personalization features, such as the ability to adjust the audio volume while using the vest and the inclusion of customizable vibration patterns and prompt frequencies. Additionally, enabling therapists to record guidance prompts in their own voices could enhance the patient experience, as one patient noted: “Listening to the app's directions reminds me of you (the therapist) talking to me during the rehabilitation sessions.”

The study’s duration might have limited the assessment of long-term effects, as improvements in posture maintenance could potentially change over time. A longer follow-up period could provide insights into whether the benefits observed are sustained in the long term. Additionally, the self-controlled approach used in the study—where participants served as their own controls by using the app with feedback turned off—may introduce bias related to user expectations and behavior. Future research should consider alternative control group designs to minimize potential biases and enhance the validity of the results. Another limitation of this study is the relatively unstructured approach to coding and scoring responses from interviews and feedback sessions. The multidimensional nature of the questions, ranging from levels of agreement to broader attitudes and preferences, presented challenges in achieving uniform analysis. To address this, we plan to use more structured, software-assisted procedures in future research to standardize data collection and analysis [68,69], thereby improving the consistency and depth of understanding gained from participant feedback.

Future research should also explore categorizing participants based on injury time, dominance, or body laterality to better understand the varying impacts of the Postural SmartVest. Including studies that involve patients using the device at home could provide insights into its real-world effectiveness and long-term impact in poststroke rehabilitation outside of clinical settings. Understanding how patients integrate the Postural SmartVest into their daily routines and the effects of extended, unsupervised use could reveal valuable information about its benefits in home-based rehabilitation. Furthermore, while Postural SmartVest records a range of activities, including posture changes and feedback delivery, it currently lacks features for therapists to access and use this data effectively. Future research should focus on developing these functionalities to enhance patient monitoring and rehabilitation planning. Additionally, addressing the requirements identified during initial therapist interviews, such as including a photo and video gallery and a dashboard for progress monitoring, could improve Postural SmartVest’s functionality and user satisfaction.

Comparison With Prior Work

Postural SmartVest is a modified athletic compression tank top with a transparent pocket for a smartphone running a customizable Android app. This app offers tactile vibrations, dynamic visual feedback, and clear audio instructions for posture monitoring and guidance. It is designed to be minimally intrusive, replicable, and useful for both therapists and poststroke patients, using readily available Android smartphones and avoiding the need for expensive specialized equipment.

The development of Postural SmartVest reflects the evolution of cognitive rehabilitation from classical methods to personalized systems with kinematic analysis and user-centered design [70]. It incorporates feedback from patients with stroke and therapists, aligning with recommendations for wearable technology design [71]. Unlike wearable rehabilitation systems that prioritize sensor accuracy over comfort [24], Postural SmartVest is distinguished by its comfort and effective smartphone placement, addressing both functional and user experience aspects.

Research on user perspectives for wearable systems in upper extremity stroke rehabilitation highlights the importance of monitoring both upper extremity and trunk movements and evaluating their quality and quantity [39]. In our experimental sessions, we included activities such as position maintenance, lateral trunk movements, trunk rotation, forward-backward trunk motions, and object retrieval, aligning with tasks assessed by various trunk ability evaluation tools [72]. While some studies have investigated smartphone-based training to improve balance and trunk performance among seated patients with stroke—using modified balance boards with visual and audio feedback in an external monitor [48,49] or smartphones attached to harnesses for screen feedback [25]—Postural SmartVest supports dynamic ambulatory movements and provides visual feedback directly on the smartphone screen, using a single color (red or green) to indicate posture status. This approach is particularly effective in therapy rooms with mirror walls and accommodates individuals with visual deficiencies, which are common after a stroke [73].

Postural SmartVest uses the smartphone’s accelerometer on the patient’s trunk to assist with trunk posture, while many studies use sensors on different body parts for various purposes [74-77]. While postural SmartVest uses smartphone-generated sensor data to identify trunk postures and facilitate posture correction, existing literature often combines data from multiple sensors placed on the user’s body for posture detection [78].

A wealth of research envisions using their solutions to foster collaboration among health professionals [79] or in the patient’s home environment, including studies that use smartphones as a communication component in a remote rehabilitation platform [18,25,33,35,70,76,80,81]. Although we designed Postural SmartVest primarily to assist patients during rehabilitation sessions, we are actively extending our research to the home environment by integrating Postural SmartVest with our experience sampling and programmed intervention platform [68].

Conclusions

Our goal was to create an affordable wearable, Postural SmartVest, aiding poststroke patients in posture maintenance during rehabilitation. Its simplicity and cost-effectiveness, using widely available devices with efficient sensors, enable multisensory feedback, enhancing accessibility. Patients and therapists expressed utility and satisfaction, emphasizing its clinical potential.

A comparison of patient session data using Postural SmartVest, both with and without feedback, informed essential insights. The significant difference in maintaining correct positions underscored the value of feedback. Although there was no disparity in the number of movements, the increased adjustments during sessions with feedback suggested active patient engagement. These findings emphasize the potential of Postural SmartVest as a valuable assistive technology for poststroke patients and highlight the importance of integrating feedback into rehabilitation tools.

Our study indicates several future research directions, including studying diverse stroke survivor populations, assessing long-term effects, exploring user engagement factors, and evaluating therapist training impact. Additionally, we should prioritize functional outcomes, technology adoption, cost-effectiveness, and addressing unmet needs like enhancing posture visualization and aiding long-term posture monitoring.

Acknowledgments

OJMN received support from the São Paulo Research Foundation (FAPESP; grant 14/01831-6). MDGCP was funded by FAPESP (grants 16/00351-6 and 16/50489-4) and CNPq (National Council for Scientific and Technological Development; grant 312058/2015-2). The authors also acknowledge the support of FAPESP for funding the publication of this manuscript (grant no. 24/15881-7). The Coordination for the Improvement of Higher Education Personnel (CAPES) contributed to this study (finance code 001). We deeply thank all participants in the interviews, feasibility study, and within-subject experiment for their participation. We thank the support from the Innovation and Entrepreneurship Laboratory in Assistive Technology, Sports, and Health at the Federal University of São Carlos for the Vicon evaluation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File , 499 KB - aging_v8i1e55455_app1.docx]

References

1. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Rajati F, Rajati M, Rasulehvandi R, Kazeminia M. Prevalence of stroke in the elderly: a systematic review and meta-analysis. Interdiscip Neurosurg 2023;32:101746. [doi: 10.1016/j.inat.2023.101746]
3. Miranda M. Números do AVC. URL: https://avc.org.br/numeros-do-avc/ [accessed 2024-12-07]
4. Hsieh CL, Sheu CF, Hsueh IP, Wang CH. Trunk control as an early predictor of comprehensive activities of daily living function in stroke patients. Stroke 2002;33(11):2626-2630. [doi: 10.1161/01.str.0000033930.05931.93] [Medline: 12411652]
5. Lin SH, Dionne TP. Interventions to improve movement and functional outcomes in adult stroke rehabilitation: review and evidence summary. J Particip Med 2018;10(1):e3 [FREE Full text] [doi: 10.2196/jopm.8929] [Medline: 33052128]
6. Ryerson S, Byl NN, Brown DA, Wong RA, Hidler JM. Altered trunk position sense and its relation to balance functions in people post-stroke. J Neurol Phys Ther 2008;32(1):14-20. [doi: 10.1097/NPT.0b013e3181660f0c] [Medline: 18463551]
7. Pellicciari L, Sodero A, Campagnini S, Guolo E, Basagni B, Castagnoli C, et al. Factors influencing trunk control recovery after intensive rehabilitation in post-stroke patients: a multicentre prospective study. Top Stroke Rehabil 2023;30(2):109-118. [doi: 10.1080/10749357.2021.2016099] [Medline: 34994302]
8. Wang C, Chen Y, Wang C. Postural maintenance is associated with walking ability in people receiving acute rehabilitation after a stroke. Phys Ther 2022;102(4):pzab309. [doi: 10.1093/ptj/pzab309] [Medline: 35079798]
9. Yang S, Chung E, Lee J, Lee S, Lee B. The effect of trunk stability training based on visual feedback on trunk stability, balance, and upper limb function in stroke patients: a randomized control trial. Healthcare (Basel) 2021;9(5):532 [FREE Full text] [doi: 10.3390/healthcare9050532] [Medline: 34063280]
10. Valdés BA, Schneider AN, Van der Loos HFM. Reducing trunk compensation in stroke survivors: a randomized crossover trial comparing visual and force feedback modalities. Arch Phys Med Rehabil 2017;98(10):1932-1940. [doi: 10.1016/j.apmr.2017.03.034] [Medline: 28526482]

11. Appleby E, Gill ST, Hayes LK, Walker TL, Walsh M, Kumar S. Effectiveness of telerehabilitation in the management of adults with stroke: a systematic review. PLoS One 2019;14(11):e0225150. [doi: 10.1371/journal.pone.0225150] [Medline: 31714924]
12. Chan A, Cohen R, Robinson K, Bhardwaj D, Gregson G, Jutai JW, et al. Evidence and user considerations of home health monitoring for older adults: scoping review. JMIR Aging 2022;5(4):e40079. [doi: 10.2196/40079] [Medline: 36441572]
13. Muroi D, Saito Y, Koyake A, Hiroi Y, Higuchi T. Training for walking through an opening improves collision avoidance behavior in subacute patients with stroke: a randomized controlled trial. Disabil Rehabil 2023;46(5):1-9. [doi: 10.1080/09638288.2023.2181412] [Medline: 36815267]
14. Sánchez-Sánchez ML, Ruescas-Nicolau M, Arnal-Gómez A, Iosa M, Pérez-Alenda S, Cortés-Amador S. Validity of an android device for assessing mobility in people with chronic stroke and hemiparesis: a cross-sectional study. J Neuroeng Rehabil 2024;21(1):54 [FREE Full text] [doi: 10.1186/s12984-024-01346-5] [Medline: 38616288]
15. Van Criekinge T, Truijen S, Schröder J, Maebe Z, Blanckaert K, van der Waal C, et al. The effectiveness of trunk training on trunk control, sitting and standing balance and mobility post-stroke: a systematic review and meta-analysis. Clin Rehabil 2019;33(6):992-1002. [doi: 10.1177/0269215519830159] [Medline: 30791703]
16. Bonura A, Motolese F, Capone F, Iaccarino G, Alessiani M, Ferrante M, et al. Smartphone app in stroke management: a narrative updated review. J Stroke 2023;25(2):320-324 [FREE Full text] [doi: 10.5853/jos.2022.01410.e1] [Medline: 37282380]
17. Chiew TK, Khoo S, Ansari P, Kiruthika U. Mobile and wearable technologies for persons with disabilities: a bibliometric analysis (2000-2021). Disabil Rehabil Assist Technol 2024;19(3):994-1002. [doi: 10.1080/17483107.2022.2138996] [Medline: 36306172]
18. Lee BC, An J, Kim J, Lai EC. Performing dynamic weight-shifting balance exercises with a smartphone-based wearable telerehabilitation system for home use by individuals with parkinson's disease: a proof-of-concept study. IEEE Trans Neural Syst Rehabil Eng 2023;31:456-463 [FREE Full text] [doi: 10.1109/TNSRE.2022.3226368] [Medline: 36455080]
19. Nascimento LMSD, Bonfati LV, Freitas MLB, Mendes Junior JJA, Siqueira HV, Stevan SL. Sensors and systems for physical rehabilitation and health monitoring-a review. Sensors (Basel) 2020;20(15):4063 [FREE Full text] [doi: 10.3390/s20154063] [Medline: 32707749]
20. Rodgers MM, Alon G, Pai VM, Conroy RS. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. J Rehabil Assist Technol Eng 2019;6:2055668319839607 [FREE Full text] [doi: 10.1177/2055668319839607] [Medline: 31245033]
21. Ding Z, Luo Z, Causo A, Chen I, Yue K, Yeo S, et al. Inertia sensor-based guidance system for upperlimb posture correction. Med Eng Phys 2013;35(2):269-276. [doi: 10.1016/j.medengphy.2011.09.002] [Medline: 21978912]
22. Lu L, Zhang J, Xie Y, Gao F, Xu S, Wu X, et al. Wearable health devices in health care: narrative systematic review. JMIR Mhealth Uhealth 2020;8(11):e18907 [FREE Full text] [doi: 10.2196/18907] [Medline: 33164904]
23. Kim GJ, Parnandi A, Eva S, Schambra H. The use of wearable sensors to assess and treat the upper extremity after stroke: a scoping review. Disabil Rehabil 2022;44(20):6119-6138 [FREE Full text] [doi: 10.1080/09638288.2021.1957027] [Medline: 34328803]
24. Wang Q, Markopoulos P, Yu B, Chen W, Timmermans A. Interactive wearable systems for upper body rehabilitation: a systematic review. J Neuroeng Rehabil 2017;14(1):20 [FREE Full text] [doi: 10.1186/s12984-017-0229-y] [Medline: 28284228]
25. Aphiphaksakul P, Siriphorn A. Home-based exercise using balance disc and smartphone inclinometer application improves balance and activity of daily living in individuals with stroke: a randomized controlled trial. PLoS One 2022;17(11):e0277870 [FREE Full text] [doi: 10.1371/journal.pone.0277870] [Medline: 36409753]
26. Hou Y, Chiu Y, Chiang S, Chen H, Sung W. Development of a smartphone-based balance assessment system for subjects with stroke. Sensors (Basel) 2019;20(1):88 [FREE Full text] [doi: 10.3390/s20010088] [Medline: 31877843]
27. Afzal MR, Byun H, Oh M, Yoon J. Effects of kinesthetic haptic feedback on standing stability of young healthy subjects and stroke patients. J Neuroeng Rehabil 2015;12:27 [FREE Full text] [doi: 10.1186/s12984-015-0020-x] [Medline: 25889581]
28. Shirvani H, Salesi M, Samadi M, Shamsoddini A. Design and development of a 3-axis accelerometer biofeedback system for real-time correction of neck posture for long-time computer usersdesign and development of a 3-axis accelerometer biofeedback system for real-time correction of neck posture for long-time computer users. J Med Signals Sens 2021;11(4):269-273 [FREE Full text] [doi: 10.4103/jmss.JMSS_56_20] [Medline: 34820299]
29. Zampogna A, Mileti I, Palermo E, Celletti C, Paoloni M, Manoni A, et al. Fifteen years of wireless sensors for balance assessment in neurological disorders. Sensors (Basel) 2020;20(11):3247 [FREE Full text] [doi: 10.3390/s20113247] [Medline: 32517315]
30. Peters J, Abou L, Wong E, Dossou MS, Sosnoff JJ, Rice LA. Smartphone-based gait and balance assessment in survivors of stroke: a systematic review. Disabil Rehabil Assist Technol 2024;19(1):177-187. [doi: 10.1080/17483107.2022.2072527] [Medline: 35584288]
31. Hou YR, Chiu YL, Chiang SL, Chen HY, Sung WH. Feasibility of a smartphone-based balance assessment system for subjects with chronic stroke. Comput Methods Programs Biomed 2018;161:191-195. [doi: 10.1016/j.cmpb.2018.04.027] [Medline: 29852961]

32. Katzan I, Schuster A, Kinzy T. Physical activity monitoring using a fitbit device in ischemic stroke patients: prospective cohort feasibility study. JMIR Mhealth Uhealth 2021;9(1):e14494 [FREE Full text] [doi: 10.2196/14494] [Medline: 33464213]
33. Wu JM, Chen HS, Chen HH, Cheng BW, Huang CW, Chung MH. Enhancing patient self-management after a first stroke: an application of the wearable devices and the health management platform. Disabil Health J 2023;16(1):101392 [FREE Full text] [doi: 10.1016/j.dhjo.2022.101392] [Medline: 36333265]
34. Seo M, Shin M, Park TS, Park J. Clinometric gait analysis using smart insoles in patients with hemiplegia after stroke: pilot study. JMIR Mhealth Uhealth 2020;8(9):e22208 [FREE Full text] [doi: 10.2196/22208] [Medline: 32909949]
35. Chae SH, Kim Y, Lee K, Park H. Development and clinical evaluation of a web-based upper limb home rehabilitation system using a smartwatch and machine learning model for chronic stroke survivors: prospective comparative study. JMIR Mhealth Uhealth 2020;8(7):e17216 [FREE Full text] [doi: 10.2196/17216] [Medline: 32480361]
36. Signal NEJ, McLaren R, Rashid U, Vandal A, King M, Almesfer F, et al. Haptic nudges increase affected upper limb movement during inpatient stroke rehabilitation: multiple-period randomized crossover study. JMIR Mhealth Uhealth 2020;8(7):e17036 [FREE Full text] [doi: 10.2196/17036] [Medline: 32723718]
37. Georgiou T, Holland S, van der Linden J. Rhythmic haptic cueing for gait rehabilitation of people with hemiparesis: quantitative gait study. JMIR Biomed Eng 2020;5(1):e18649. [doi: 10.2196/18649]
38. Shim E, Lee S, Lee H, Choi YS. Smart pose: mobile posture-aware system for lowering physical health risk of smartphone users. : Association for Computing Machinery; 2013 Presented at: CHI '13 Extended Abstracts on Human Factors in Computing Systems; 2013 April 27; Paris, France p. 2257-2266 URL: https://doi.org/10.1145/2468356.2468747
39. Wang Q. Designing posture monitoring garments to support rehabilitation. 2016 Presented at: TEI '16: Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction; 2016 February 14; Eindhoven, Netherlands p. 709-712 URL: https://dl.acm.org/doi/10.1145/2839462.2854106
40. Ribeiro P, Soares AR, Girão R, Neto M, Cardoso S. Spine cop: posture correction monitor and assistant. Sensors (Basel) 2020;20(18):5376 [FREE Full text] [doi: 10.3390/s20185376] [Medline: 32961772]
41. Muşat EC, Borz SA. Learning from acceleration data to differentiate the posture, dynamic and static work of the back: an experimental setup. Healthcare (Basel) 2022;10(5):916 [FREE Full text] [doi: 10.3390/healthcare10050916] [Medline: 35628053]
42. Yang CL, Chui R, Mortenson WB, Servati P, Servati A, Tashakori A, et al. Perspectives of users for a future interactive wearable system for upper extremity rehabilitation following stroke: a qualitative study. J Neuroeng Rehabil 2023;20(1):77 [FREE Full text] [doi: 10.1186/s12984-023-01197-6] [Medline: 37312189]
43. Soares CL, Magalhães JP, Faria-Fortini I, Batista LR, Andrea Oliveira Lima L, Faria CD. Barriers and facilitators to access post-stroke rehabilitation services in the first six months of recovery in Brazil. Disabil Rehabil 2024;46(23):5601-5607. [doi: 10.1080/09638288.2024.2310756] [Medline: 38299553]
44. Nepal G, Yadav JK, Bhandari S, Gautam J, Gajurel BP. Low-cost alternatives for the management of acute ischemic stroke in low and middle-income countries. Ann Med Surg (Lond) 2021;72:102969 [FREE Full text] [doi: 10.1016/j.amsu.2021.102969] [Medline: 34992776]
45. Ariza J, Pearce JM. Low-cost assistive technologies for disabled people using open-source hardware and software: a systematic literature review. IEEE Access 2022;10:124894-124927. [doi: 10.1109/access.2022.3221449]
46. Negrini F, Gasperini G, Guanziroli E, Vitale JA, Banfi G, Molteni F. Using an accelerometer-based step counter in post-stroke patients: validation of a low-cost tool. Int J Environ Res Public Health 2020;17(9):3177 [FREE Full text] [doi: 10.3390/ijerph17093177] [Medline: 32370210]
47. Burns SP, Terblanche M, Perea J, Lillard H, DeLaPena C, Grinage N, et al. mHealth intervention applications for adults living with the effects of stroke: a scoping review. Arch Rehabil Res Clin Transl 2021;3(1):100095 [FREE Full text] [doi: 10.1016/j.arrct.2020.100095] [Medline: 33778470]
48. Shin DC, Song CH. Smartphone-based visual feedback trunk control training using a gyroscope and mirroring technology for stroke patients: single-blinded, randomized clinical trial of efficacy and feasibility. Am J Phys Med Rehabil 2016;95(5):319-329. [doi: 10.1097/PHM.0000000000000447] [Medline: 26829087]
49. Chung BPH, Chiang WKH, Lau H, Lau TFO, Lai CWK, Sit CSY, et al. Pilot study on comparisons between the effectiveness of mobile video-guided and paper-based home exercise programs on improving exercise adherence, self-efficacy for exercise and functional outcomes of patients with stroke with 3-month follow-up: a single-blind randomized controlled trial. Hong Kong Physiother J 2020;40(1):63-73 [FREE Full text] [doi: 10.1142/S1013702520500079] [Medline: 32489241]
50. van Vliet PM, Wulf G. Extrinsic feedback for motor learning after stroke: what is the evidence? Disabil Rehabil 2006;28(13-14):831-840. [doi: 10.1080/09638280500534937] [Medline: 16777770]
51. Soares M, Bonuzzi GMG, Coelho DB, Torriani-Pasin C. Effects of extrinsic feedback on the motor learning after stroke. Motriz rev educ fis 2019;25(1). [doi: 10.1590/s1980-6574201900010013]
52. Carr K, Zachariah N, Weir P, McNevin N. An examination of feedback use in rehabilitation settings. Crit Rev Phys Rehabil Med 2011;23(1-4):147-160. [doi: 10.1615/critrevphysrehabilmed.v23.i1-4.110]
53. Stanton R, Ada L, Dean CM, Preston E. Feedback received while practicing everyday activities during rehabilitation after stroke: an observational study. Physiother Res Int 2015;20(3):166-173. [doi: 10.1002/pri.1612] [Medline: 25451759]

54. Johnson L, Burridge J, Demain S. Internal and external focus of attention during gait re-education: an observational study of physical therapist practice in stroke rehabilitation. Phys Ther 2013;93(7):957-966. [doi: 10.2522/ptj.20120300] [Medline: 23559523]
55. Figueira V, Silva S, Costa I, Campos B, Salgado J, Pinho L, et al. Wearables for monitoring and postural feedback in the work context: a scoping review. Sensors (Basel) 2024;24(4):1341 [FREE Full text] [doi: 10.3390/s24041341] [Medline: 38400498]
56. Machado Neto OJ, Peracini APP, Cunha BCR, Rocha ALC, Elui VMC, Castro PC, et al. An architecture for creating mobile assistive products: a case study on posture monitoring of people with motor disability. 2016 Presented at: 2nd IET International Conference on Technologies for Active and Assisted Living (TechAAL 2016); 2016 October 24-25; London, UK URL: https://ieeexplore.ieee.org/document/7801339
57. Ometov A, Shubina V, Klus L, Skibińska J, Saafi S, Pascacio P, et al. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. Comput Networks 2021 Jul;193:108074. [doi: 10.1016/j.comnet.2021.108074]
58. Heebner NR, Akins JS, Lephart SM, Sell TC. Reliability and validity of an accelerometry based measure of static and dynamic postural stability in healthy and active individuals. Gait Posture 2015;41(2):535-539. [doi: 10.1016/j.gaitpost.2014.12.009] [Medline: 25544692]
59. Guest G, MacQueen K, Namey E. Applied Thematic Analysis. Washington, DC: Sage Publications; 2014.
60. Demers L, Weiss-Lambrou R, Ska B. The Quebec User Evaluation of Satisfaction with Assistive Technology (QUEST 2.0): an overview and recent progress. Technol Disabil 2002;14(3):101-105. [doi: 10.3233/tad-2002-14304]
61. de Carvalho KEC, Gois Júnior MB, Sá KN. Translation and validation of the Quebec User Evaluation of Satisfaction With Assistive Technology (QUEST 2.0) into Portuguese. Rev Bras Reumatol 2014;54(4):260-267 [FREE Full text] [doi: 10.1016/j.rbr.2014.04.003] [Medline: 25627220]
62. Jayaraman C, Mummidisetty CK, Mannix-Slobig A, McGee Koch L, Jayaraman A. Variables influencing wearable sensor outcome estimates in individuals with stroke and incomplete spinal cord injury: a pilot investigation validating two research grade sensors. J Neuroeng Rehabil 2018;15(1):19 [FREE Full text] [doi: 10.1186/s12984-018-0358-y] [Medline: 29534737]
63. Grouios G, Ziagkas E, Loukovitis A, Chatzinikolaou K, Koidou E. Accelerometers in our pocket: does smartphone accelerometer technology provide accurate data? Sensors (Basel) 2022;23(1):192 [FREE Full text] [doi: 10.3390/s23010192] [Medline: 36616798]
64. Riberto M, Miyazaki MH, Jucá SSH, Sakamoto H, Potiguara P. Validação da Versão Brasileira da Medida de Independência Funcional Validation of the Brazilian version of Functional Independence Measure. Acta Fisiatrica 2004;11 [FREE Full text]
65. Kidd D, Stewart G, Baldry J, Johnson J, Rossiter D, Petruckevitch A, et al. The functional independence measure: a comparative validity and reliability study. Disabil Rehabil 1995;17(1):10-14. [doi: 10.3109/09638289509166622] [Medline: 7858276]
66. Pedrosa D, Pimentel MDGC, Wright A, Truong KN. Filteryedping: design challenges and user performance of dwell-free eye typing. ACM Trans Access Comput 2015;6(1):1-37. [doi: 10.1145/2724728]
67. Open mHealth. 2019. URL: https://www.openmhealth.org/ [accessed 2024-11-30]
68. Cunha BCR, Rodrigues KRDH, Zaine I, da Silva EAN, Viel CC, Pimentel MDGC. Experience sampling and programmed intervention method and system for planning, authoring, and deploying mobile health interventions: design and case reports. J Med Internet Res 2021;23(7):e24278 [FREE Full text] [doi: 10.2196/24278] [Medline: 34255652]
69. Dantas L, Pimentel MDGC, Melo R, De Lima AP. Analysis of a digital instrument for multidimensional assessment of the older adults by undergraduate gerontology students. Cuad Ed Desar 2023;15(9):9168-9185.
70. Cogollor JM, Rojo-Lacal J, Hermsdörfer J, Ferre M, Arredondo Waldmeyer MT, Giachritsis C, et al. Evolution of cognitive rehabilitation after stroke from traditional techniques to smart and personalized home-based information and communication technology systems: literature review. JMIR Rehabil Assist Technol 2018;5(1):e4 [FREE Full text] [doi: 10.2196/rehab.8548] [Medline: 29581093]
71. Sorrentino G, Sale P, Solaro C, Rabini A, Cerri CG, Ferriero G. Clinical measurement tools to assess trunk performance after stroke: a systematic review. Eur J Phys Rehabil Med 2018;54(5):772-784 [FREE Full text] [doi: 10.23736/S1973-9087.18.05178-X] [Medline: 29684980]
72. Elnady A, Mortenson WB, Menon C. Perceptions of existing wearable robotic devices for upper extremity and suggestions for their development: findings from therapists and people with stroke. JMIR Rehabil Assist Technol 2018;5(1):e12 [FREE Full text] [doi: 10.2196/rehab.9535] [Medline: 29764799]
73. Hazelton C, McGill K, Campbell P, Todhunter-Brown A, Thomson K, Nicolson DJ, et al. Perceptual disorders after stroke: a scoping review of interventions. Stroke 2022;53(5):1772-1787 [FREE Full text] [doi: 10.1161/STROKEAHA.121.035671] [Medline: 35468001]
74. Huo C, Zheng Y, Lu W, Zhang T, Wang D, Xu D, et al. Prospects for intelligent rehabilitation techniques to treat motor dysfunction. Neural Regen Res 2021;16(2):264-269 [FREE Full text] [doi: 10.4103/1673-5374.290884] [Medline: 32859773]
75. Guo L, Wang J, Wu Q, Li X, Zhang B, Zhou L, et al. Clinical study of a wearable remote rehabilitation training system for patients with stroke: randomized controlled pilot trial. JMIR Mhealth Uhealth 2023;11:e40416 [FREE Full text] [doi: 10.2196/40416] [Medline: 36821348]

76. Toh SFM, Fong KNK, Gonzalez PC, Tang YM. Application of home-based wearable technologies in physical rehabilitation for stroke: a scoping review. IEEE Trans Neural Syst Rehabil Eng 2023. [doi: 10.1109/TNSRE.2023.3252880] [Medline: 37028029]
77. Davies RJ, Parker J, McCullagh P, Zheng H, Nugent C, Black ND, et al. A personalized self-management rehabilitation system for stroke survivors: a quantitative gait analysis using a smart insole. JMIR Rehabil Assist Technol 2016;3(2):e11 [FREE Full text] [doi: 10.2196/rehab.5449] [Medline: 28582260]
78. Fonte G, Schreiber C, Areno G, Masson X, Chantraine F, Schütz G, et al. Metabolic energy expenditure and accelerometer-determined physical activity levels in post-stroke hemiparetic patients. J Stroke Cerebrovasc Dis 2022;31(5):106397. [doi: 10.1016/j.jstrokecerebrovasdis.2022.106397] [Medline: 35228022]
79. Ploderer B, Fong J, Klaic M, Nair S, Vetere F, Cofré Lizama LE, et al. How therapists use visualizations of upper limb movement information from stroke patients: a qualitative study with simulated information. JMIR Rehabil Assist Technol 2016;3(2):e9 [FREE Full text] [doi: 10.2196/rehab.6182] [Medline: 28582257]
80. Toh SFM, Gonzalez PC, Fong KNK. Usability of a wearable device for home-based upper limb telerehabilitation in persons with stroke: a mixed-methods study. Digit Health 2023;9:20552076231153737 [FREE Full text] [doi: 10.1177/20552076231153737] [Medline: 36776407]
81. Lin WTM, Lin BS, Lee IJ, Lee SH. Development of a smartphone-based mHealth platform for telerehabilitation. IEEE Trans Neural Syst Rehabil Eng 2022;30:2682-2691. [doi: 10.1109/TNSRE.2022.3204148] [Medline: 36063516]

Abbreviations

QUEST: Quebec User Evaluation of Satisfaction with Assistive Technology

Edited by Q Fan; submitted 13.12.23; peer-reviewed by C Abbatantuono; comments to author 01.04.24; revised version received 29.08.24; accepted 06.09.24; published 22.01.25.

Please cite as:
Pereira AP, Machado Neto OJ, Elui VMC, Pimentel MDGC
Wearable Smartphone-Based Multisensory Feedback System for Torso Posture Correction: Iterative Design and Within-Subjects Study
JMIR Aging 2025;8:e55455
URL: https://aging.jmir.org/2025/1/e55455
doi:10.2196/55455
PMID:

Original Paper

Estimation of Machine Learning–Based Models to Predict Dementia Risk in Patients With Atherosclerotic Cardiovascular Diseases: UK Biobank Study

Zhengsheng Gu[1*], MD; Shuang Liu[2*], MM; Huijuan Ma[2*], PhD; Yifan Long[2], MM; Xuehao Jiao[1], MD; Xin Gao[1], MM; Bingying Du[1,3*], PhD; Xiaoying Bi[1*], PhD; Xingjie Shi[2*], PhD

[1]Department of Neurology, First Affiliated Hospital of Naval Medical University, Shanghai, China

[2]KLATASDS-MOE, Academy of Statistics and Interdisciplinary Sciences, School of Statistics, East China Normal University, Shanghai, China

[3]Institute for Translational Brain Research, Fudan University, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Xingjie Shi, PhD
KLATASDS-MOE
Academy of Statistics and Interdisciplinary Sciences, School of Statistics
East China Normal University
No 3663 North Zhongshan Road, Putuo District
Shanghai, 200062
China
Phone: 86 21 622 332 23
Email: xjshi@fem.ecnu.edu.cn

Abstract

Background: The atherosclerotic cardiovascular disease (ASCVD) is associated with dementia. However, the risk factors of dementia in patients with ASCVD remain unclear, necessitating the development of accurate prediction models.

Objective: The aim of the study is to develop a machine learning model for use in patients with ASCVD to predict dementia risk using available clinical and sociodemographic data.

Methods: This prognostic study included patients with ASCVD between 2006 and 2010, with registration of follow-up data ending on April 2023 based on the UK Biobank. We implemented a data-driven strategy, identifying predictors from 316 variables and developing a machine learning model to predict the risk of incident dementia, Alzheimer disease, and vascular dementia within 5, 10, and longer-term follow-up in patients with ASCVD.

Results: A total of 29,561 patients with ASCVD were included, and 1334 (4.51%) developed dementia during a median follow-up time of 10.3 (IQR 7.6-12.4) years. The best prediction model (UK Biobank ASCVD risk prediction model) was light gradient boosting machine, comprising 10 predictors including age, time to complete pairs matching tasks, mean time to correctly identify matches, mean sphered cell volume, glucose levels, forced expiratory volume in 1 second *z* score, C-reactive protein, forced vital capacity, time engaging in activities, and age first had sexual intercourse. This model achieved the following performance metrics for all incident dementia: area under the receiver operating characteristic curve: mean 0.866 (SD 0.027), accuracy: mean 0.883 (SD 0.010), sensitivity: mean 0.637 (SD 0.084), specificity: mean 0.914 (SD 0.012), precision: mean 0.479 (SD 0.031), and F_1-score: mean 0.546 (SD 0.043). Meanwhile, this model was well-calibrated (Kolmogorov-Smirnov test showed goodness-of-fit *P* value>.99) and maintained robust performance across different temporal cohorts. Besides, the model had a beneficial potential in clinical practice with a decision curve analysis.

Conclusions: The findings of this study suggest that predictive modeling could inform patients and clinicians about ASCVD at risk for dementia.

(JMIR Aging 2025;8:e64148) doi:10.2196/64148

KEYWORDS

atherosclerotic cardiovascular disease; dementia; Alzheimer disease; vascular dementia; machine learning; UK Biobank

Introduction

Cardiovascular disease (CVD) is the leading cause of noncommunicable disease and mortality worldwide [1]. Meanwhile, the epidemiology of the atherosclerotic cardiovascular disease (ASCVD), which encompasses coronary heart disease and cerebrovascular disease (CeVD), has experienced substantial and rapid growth [2]. It is reported that in 2016, ASCVD was responsible for approximately 2.4 million deaths, representing 25% of all deaths and 61% of CVD-related deaths in China [3].

Dementia is another devastating disease affecting more than 50 million individuals worldwide [4]. Given the high costs and heavy burdens it imposes on families and society, scientists and scholars around the world are dedicated to identifying preventable interventions and reducing the incidence of dementia. Recently, a growing body of evidence indicates that lifestyle interventions early in life with a focus on reducing cardiovascular risk factors are a promising strategy for preventing dementia [5-9]. In particular, shared risk factors between dementia and ASCVD have been identified [10]. According to the Lancet Commission, it is estimated that approximately 40% of dementia cases can be prevented by targeting modifiable, primarily cardiovascular risk factors [4]. However, these studies were restricted by their use of classical statistical analyses (such as Cox or logistic regressions) and by considering only widely studied prespecified CVD risks. Therefore, the results were not sufficient in accuracy.

Machine learning (ML) is an emerging technical foundation of artificial intelligence, which enables the leverage of information from large and complex datasets [11]. Several studies have applied ML-based models to dementia diagnosis and risk prediction [12-15]. Nevertheless, the long-term risk of dementia progression (5 or 10 years) in patients with ASCVD remains uncertain.

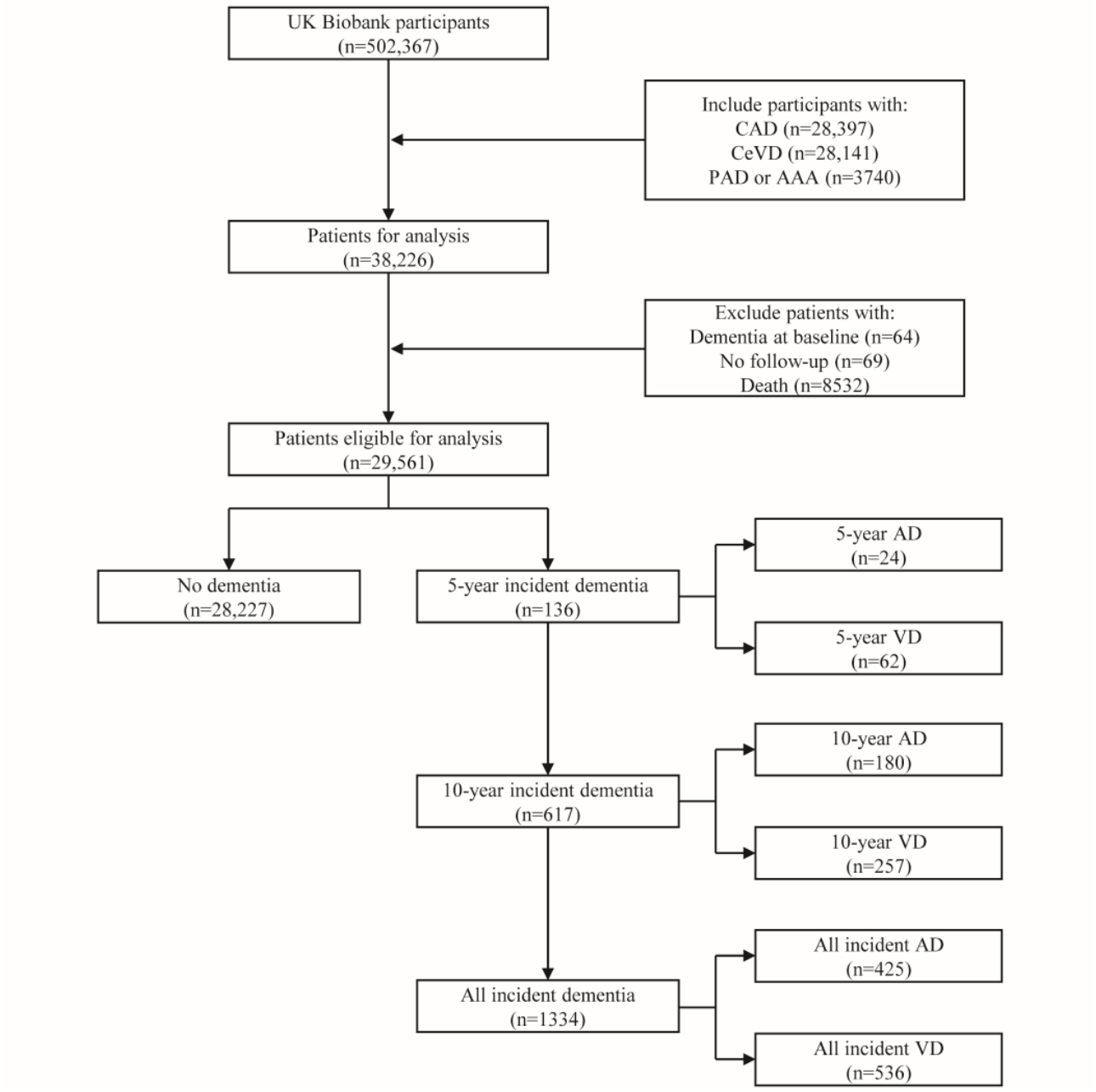
In this study, we used comprehensive phenotypic and follow-up data from a cohort of over 500,000 UK Biobank participants to develop an ML-based model capable of predicting the 5- or 10-year risk of incident dementia in specific patients with ASCVD. We anticipate that this ML-derived early warning system will enhance clinician-patient counseling, enable targeted follow-up, and facilitate the development of personalized prevention strategies. This, ultimately, can optimize the health and care of individuals with ASCVD.

Methods

Data Source and Study Population

We analyzed data from the UK Biobank, a longitudinal prospective study that recruited over 500,000 participants between 2006 and 2010 [16]. The participants were enrolled from 22 recruitment centers across the United Kingdom and were aged between 40 and 69 years at the baseline assessment. Multiple data were collected from the participants, including questionnaires, physical measurements, biological sample assays, genotyping, imaging data, and ongoing hospital records. Figure 1 illustrated the enrollment process, where we included individuals with a prior history of any established ASCVD, such as coronary artery disease (n=28,397), CeVD (n=28,141), peripheral artery disease, or abdominal aortic aneurysm (n=3740) [17]. Participants were excluded at the baseline assessment if they met the following criteria: (1) had dementia at baseline (n=64), (2) had no follow-up records (n=69), and (3) death (n=8532). Ultimately, we included 29,561 participants with ASCVD who had at least 10 years of follow-up until April 2023.

Figure 1. Participant selection flowchart. UK Biobank participants were excluded if baseline dementia was self-reported or follow-up data were absent. The remaining participants were categorized according to their first reported years of dementia, AD, or VD after the baseline. AAA: abdominal aortic aneurysm; AD: Alzheimer disease; CAD: coronary artery disease; CeVD: cerebrovascular disease; PAD: peripheral artery disease; VD: vascular dementia.



Ethical Considerations

Ethical approval was obtained from the North West Multi-Centre Research Ethics Committee (11/NW/0382, 16/NW/0274, and 21/NW/0157). Written informed consent was provided by all participants during the collection of primary data. The UK Biobank data used were deidentified, and all personally identifiable information of participants has been removed to ensure privacy and confidentiality. Besides, the UK Biobank offered nonfinancial compensation in the form of travel reimbursements for attending the assessment centers and other incidental expenses related to participation. Additionally, participants were given feedback on their individual health data upon request, which provided valuable insights into their health status. This study adhered to the reporting guidelines of Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis [18].

Outcome

The primary end point of this study was the occurrence of all incident dementia, including Alzheimer disease (AD), vascular dementia (VD), frontotemporal dementia, and dementia associated with other neurodegenerative or specified diseases. Due to the high rate of incidence worldwide, AD and VD were examined as the secondary outcomes. To conduct a comprehensive survey on the incidence time, we categorized the patients into 5-year, 10-year, and all incident dementia, AD, and VD. The outcomes were ascertained and categorized based on the International Classification of Diseases and Read codes (Table S1 in Multimedia Appendix 1), which were obtained

from the "first occurrence" category in the UK Biobank including the primary care data, the hospital inpatient data, the death register records, and subsequent UK Biobank assessment center visits. Follow-up visits continued until the earliest of the following events: a dementia diagnosis, death, or the most recent available data from either the hospital or the general practitioner, whichever occurred first. What is noteworthy was that the imaging data and lumbar puncture results were not available to doctors to achieve the detailed diagnostic data.

Data Preparation

In this study, we included all clinically correlated variables during the participants' baseline visits. The assessment procedure involved a manual examination of each variable to determine its relevance to comprehensively understanding a participant's overall status. Variables not pertinent to these key domains or lacking in additional insights were excluded. Data screening was processed to exclude noninformative variables with missing values exceeding 40% among all participants. To prevent potential overfitting from oversampling, we applied random undersampling to the majority class, balancing the dataset more effectively. We also adjusted the class weights in our ML algorithms to give more importance to the minority class during model training. Overall, a total of 316 features were adopted, including the participants' demographic characteristics (n=2), touchscreen-recorded questionnaires (n=151), physical measures (n=66), cognitive function tests (n=22), and biological sample assays (n=59). Furthermore, to improve the informative value of the dataset, we used the available data to generate several variables (n=16) that were not directly extracted from the UK Biobank (Table S2 in Multimedia Appendix 1). Considering the significant impact of ASCVD on mortality, we identified and coded deaths as competing events to ensure accurate modeling of the primary outcome.

In this study, we used different missing data handling strategies tailored to each ML algorithm to ensure the accuracy and robustness of the models. Specifically, for the logistic regression model, we used mean imputation to handle missing values. For each variable with missing data, we calculated its mean in the training dataset and replaced the missing data points with this mean. This method is simple and efficient, making it suitable for models like logistic regression that require complete datasets. For other ML methods, including light gradient boosting machine (LightGBM), extreme gradient boosting machine, random forest, k-nearest neighbor, and artificial neural network, we adopted automatic imputation techniques. These automatic imputation methods leverage the inherent mechanisms of the algorithms or advanced imputation strategies within the preprocessing pipeline to dynamically estimate and replace missing values. To evaluate the robustness of our results, we conducted a sensitivity analysis using multiple imputation by chained equations (Table S4 in Multimedia Appendix 1). This approach generates several imputed datasets by modeling each missing value conditionally based on other variables, thereby accounting for the uncertainty associated with the imputations. By comparing the outcomes across different imputation methods, we assessed the stability and reliability of our predictive models.

To evaluate the model's stability and generalization across different time periods, this study used a time validation approach to partition dementia data from the UK Biobank database. First, the recruitment date was selected as the primary temporal variable, and all samples were sorted in ascending order based on this date to ensure chronological arrangement and prevent future data from leaking into the training process. Considering previous research and data volume, the dataset was divided into 2 periods: the training and validation sets comprised samples recruited from 2006 to 2009, while the test set included samples diagnosed in 2010. This division ensures that only past data were used for model training, and the model's predictive performance on future data was assessed during the validation and testing phases. To further guarantee temporal independence, feature selection and standardization were performed exclusively on the training set, with identical transformations applied to the validation and test sets, thereby avoiding the use of information from these sets during training. Given the typically low number of dementia cases, healthy samples with more than 5% missing variables were excluded from the training set to balance the class distribution and enhance model learning. Additionally, multiple imputation methods were used to handle missing data, ensuring data integrity. Through these steps, a time validation framework was established, maintaining the temporal independence and appropriate distribution of the training, validation, and test sets, thereby improving the model's predictive performance across different time periods and the credibility of the study's findings.

Predictor Selection

The predictors for model development were identified through a 2-step process: variable importance ranking and sequential forward selection (SFS) [19,20]. First of all, the importance of each variable was determined using a preliminary trained LightGBM classifier. Gradient boosting machine is a type of boosting that builds these simple models step-by-step, improving the model at each step to better fit the data. LightGBM, developed by Microsoft, is a faster and more efficient version of gradient boosting machine designed to handle large-scale data effectively. The "light" in LightGBM refers to its lightweight nature, meaning it uses less memory and runs faster. The top 50 variables were selected by LightGBM. Next, they were inputted into a hierarchical clustering algorithm, which used Spearman rank-order correlations to further identify and eliminate redundant variables with multicollinearity. We established a correlation threshold of 0.75, considering variables with pairwise correlations above this value as highly redundant. Within each cluster of such variables, we retained only the most predictive variable for the model, effectively reducing multicollinearity while preserving essential information. To avoid overfitting and enhance the robustness of feature selection, a nested cross-validation approach was used. Specifically, in the outer loop, we divided the dataset into multiple folds, selecting 1 fold as the test set and using the remaining folds for feature selection and model training. Within the inner loop, the training set was further split into inner training and validation sets, where features were selected based on performance in the inner validation sets. Finally, model performance was evaluated on the outer test set to ensure fair feature selection and robust

predictive capability. Then, an SFS approach was used, wherein the features within the preselected subset underwent reranking according to a newly developed classifier. Afterward, preselected variables were reranked, and multiple ML classifiers were used to sequentially add predictors one at each time. Finally, the classifier was selected based on achieving the best performance of area under the receiver operating characteristic curve (AUC), and we selected the top 10 variables according to the importance of each variable calculated by the LightGBM model. After selecting these 10 variables, adding any other variables did not significantly improve the model. The top 25 predictors are shown in Table 1. More details of predictor selection could be obtained in the part of the Methods section in Multimedia Appendix 1.

Table 1. Top 25 predictors for all incident dementia with light gradient boosting machine.

Number	Variables	Importance rating	Ranking
1	Forced vital capacity	0.133	1
2	Summed MET[a] minutes per week for all activity	0.12	2
3	Age	0.09	3
4	Pairs matching time	0.05	4
5	Mean sphered cell volume	0.038	5
6	Glucose	0.037	6
7	Mean time to correctly identify matches	0.037	7
8	FEV1[b] *z* score	0.035	8
9	Age first had sexual intercourse	0.033	9
10	C-reactive protein	0.031	10
11	Average weekly red wine intake	0.029	11
12	Calcium	0.029	12
13	Vitamin D	0.028	13
14	Pulse rate automated reading	0.026	14
15	Father age at death	0.025	15
16	Systolic blood pressure automated reading array	0.025	16
17	FEV1/FVC[c] ratio *z* score	0.024	17
18	Neuroticism score	0.023	18
19	Red blood cell erythrocyte distribution width	0.023	19
20	Apolipoprotein B	0.023	20
21	Total bilirubin	0.023	21
22	Cystatin C	0.022	22
23	Alanine aminotransferase	0.018	23
24	Average weekly beer plus cider intake	0.017	24
25	Result ranking	0.016	25

[a]MET: metabolic equivalent.

[b]FEV1: forced expiratory volume in 1 second.

[c]FVC: forced vital capacity.

Model Development

We implemented a range of ML techniques, including LightGBM, extreme gradient boosting machine, random forest, logistic regression, k-nearest neighbor, support vector machine, and artificial neural network to classify participants into 2 classes: 0 (predicted to remain no dementia) or class 1 (to develop all incident dementia, AD, or VD). The proposed model was developed using patients with ASCVD without dementia (n=28,227) and with all incident dementia (n=1334) from the UK Biobank dataset. In total, 10 identified predictors were incorporated into the model. We expanded our performance evaluation metrics to include receiver operating characteristic-AUC, precision, recall, and F_1-score, ensuring a comprehensive assessment of the models' performance on imbalanced data. Subsequently, LightGBM, the best-performing method, was used to develop a dementia risk prediction model of ASCVD, named the UK Biobank ASCVD risk prediction model. The hyperparameter tuning was performed through exhaustive selection from 10,000 sets of candidate parameters,

and the optimal set was chosen based on the performance measurement of AUC. Please refer to Table S8 in Multimedia Appendix 1 for detailed information on the search space and final adopted parameters. To evaluate the predictive performance of the models, we constructed and compared the traditional Cox proportional hazards model with the LightGBM model. Both models used identical predictor variables to ensure fairness and consistency in the comparison. The Cox model assessed hazard ratios for each variable through multivariate regression analysis, while the LightGBM model leveraged its robust ability to handle nonlinear relationships and variable interactions for risk prediction. Subsequently, the performance of both models was systematically compared using consistent evaluation metrics (such as AUC) to determine their predictive effectiveness within the study dataset. This comparison aims to validate the potential advantages of the LightGBM model in risk prediction and to provide a reference for the application of the traditional Cox model. The Cox proportional hazards model was also used to account for competing risks, which ensured that the risk of death did not bias the estimation of dementia event probabilities. To enhance the model's stability and applicability, a time validation approach was used for data analysis. Additionally, we calibrated the raw predicted probabilities into actual dementia risks (Figures S12-S14 in Multimedia Appendix 1). Finally, to assess the clinical utility of the prediction model, decision curve analysis (DCA) was conducted. First, the model's net benefit was calculated across various threshold probabilities and then compared with the baseline strategies of "treat all" and "treat none." The DCA curves were plotted using the *rmda* package in R (R Foundation for Statistical Computing) to illustrate the model's potential value in clinical decision-making. The ML algorithm was implemented using LightGBM library (version 3.3.2) and scikit-learn library (version 1.0.2) in Python (version 3.9; Python Software Foundation).

We also performed a 5-fold cross-validation to assess the stability of feature importance, randomly dividing the dataset into 5 equal parts. In each iteration, 4 folds were used for training, and 1 fold for validation. The training involved 2 stages: model development and calibration. The 4 training folds were split 3:1, with 3 folds for development and 1 fold for calibration. Validation sets were exclusively for performance evaluation. Results were averaged across folds with corresponding SDs.

Statistical Analysis

In an analysis of the variables of interest, continuous variables were summarized using the median and IQR, while discrete variables were summarized using frequency and percentage. Group comparisons (no dementia vs incident dementia or AD or VD) were conducted using chi-square tests for discrete variables and 2-tailed Student *t* tests for continuous variables. Multivariate analysis was used to calculate odds ratios based on normalized data.

The model's performance was evaluated using 2 accuracy metrics: discrimination and calibration. Discrimination was assessed using the AUC, which ranges from 0.5 for a noninformative model to 1 for a perfectly discriminating model. Calibration measures the agreement between predicted probabilities and observed event proportions. It was evaluated using the Kolmogorov-Smirnov test with 10 subgroups and visually represented in calibration plots. A *P* value greater than .05 signified an adequate goodness of fit.

Furthermore, we reported accuracy, sensitivity, specificity, precision, and the F_1-score, which were determined using the cutoff that maximized the Youden index. Additionally, we used Shapley Additive Explanations (SHAP) plots to visualize the individual contributions of each predictor to the target variable. All data analysis and visualizations were performed using Python (version 3.9) with packages from the scikit-learn library (version 1.0.2) and the SHAP library (version 0.40.0).

Results

Population Characteristics

After quality control, a total of 29,561 participants with ASCVD were included in this study. The median age of the participants was 62.0 (IQR 58.0-66.0) years. Among the participants, 36.63% (10,829/29,561) were women, and 94.12% (27,822/29,561) were White. During a median follow-up time of 10.3 (IQR 7.6-12.4) years, a subset of 1334 participants developed dementia after their baseline visits. Specifically, 617 participants had incidents within 10 years, and 136 had incidents within 5 years. Besides, the prevalence of all-cause dementia was 4.51% (1334/29,561), AD was 1.44% (425/29,561), and VD was 1.81% (536/29,561) in this study. The critical baseline predictors are presented by incident dementia, AD, and VD status in Table 2, and the percentage of missing values for the predictors is shown in Table S3 in Multimedia Appendix 1.

Table 2. The baseline characteristics of UK Biobank participants included in the study by dementia, Alzheimer disease (AD), and vascular dementia (VD) status.

Participants characteristics	Overall (n=29,561)	No dementia (n=28,227)	All incident dementia (n=1334)	All incident AD (n=425)	All incident VD (n=536)
Age (years), median (IQR)	62.0 (58.0-66.0)	62.0 (57.0-66.0)	66.0 (63.0-68.0)	66.0 (64.0-68.0)	66.0 (63.0-68.0)
Sex (female), n (%)	10,829 (36.63)	10,353 (36.68)	476 (35.68)	170 (40.00)	163 (30.41)
Ethnicity (White), n (%)	27,822 (94.12)	26,565 (94.11)	1257 (94.23)	404 (95.06)	504 (94.03)
Education (years), median (IQR)	10.0 (9.0-11.0)	10.0 (9.0-11.0)	10.0 (9.0-11.0)	9.0 (9.0-10.0)	9.0 (9.0-10.0)
Forced vital capacity (L), median (IQR)	3.5 (2.9-4.2)	3.5 (2.9-4.2)	3.2 (2.6-3.9)	3.2 (2.7-4.0)	3.1 (2.6-3.9)
Summed MET[a] minutes per week for all activity (minutes per week), median (IQR)	1662.0 (693.0-3546.0)	1671.0 (698.0-3546.0)	1398.0 (510.0-3288.0)	1653.0 (660.0-3546.0)	1308.0 (408.8-2942.6)
Pairs matching time (seconds), median (IQR)	411.0 (325.0-534.0)	408.0 (324.0-529.0)	487.5 (368.0-678.8)	487.0 (372.0-673.5)	496.0 (370.5-701.0)
Mean sphered cell volume (fL), median (IQR)	82.4 (79.2-85.9)	82.4 (79.2-85.8)	83.1 (79.5-86.7)	82.5 (79.6-86.1)	83.2 (79.2-87.3)
Glucose (mmol/L), median (IQR)	5.0 (4.7-5.5)	5.0 (4.7-5.5)	5.1 (4.7-5.9)	5.1 (4.7-5.6)	5.2 (4.7-6.2)
Mean time to correctly identify matches (seconds), median (IQR)	563.0 (500.0-644.0)	562.0 (500.0-641.0)	594.0 (531.0-699.8)	586.0 (527.0-684.0)	602.0 (532.0-707.5)
FEV1[b]z score (L), median (IQR)	0.6 (–0.1 to 1.3)	0.6 (–0.1 to 1.3)	0.7 (0.0-1.5)	0.5 (–0.0 to 1.4)	0.8 (0.2-1.6)
Age first had sexual intercourse (years), median (IQR)	18.0 (16.0-21.0)	18.0 (16.0-21.0)	18.0 (17.0-21.0)	19.0 (17.0-21.0)	18.0 (16.0-21.0)
C-reactive protein (mg/L), median (IQR)	1.5 (0.8-3.1)	1.5 (0.8-3.1)	1.5 (0.7-3.2)	1.5 (0.7-3.3)	1.7 (0.8-3.4)
CAD[c], n (%)	21,735 (73.53)	20,763 (73.56)	972 (72.86)	325 (76.47)	357 (66.60)
CeVD[d], n (%)	21,707 (73.43)	20,738 (73.47)	969 (72.64)	324 (76.24)	356 (66.42)
PAD[e] or AAA[f], n (%)	1345 (4.55)	1270 (4.50)	75 (5.62)	23 (5.41)	34 (6.34)

[a]MET: metabolic equivalent.

[b]FEV1: forced expiratory volume in 1 second.

[c]CAD: coronary artery disease.

[d]CeVD: cerebrovascular disease.

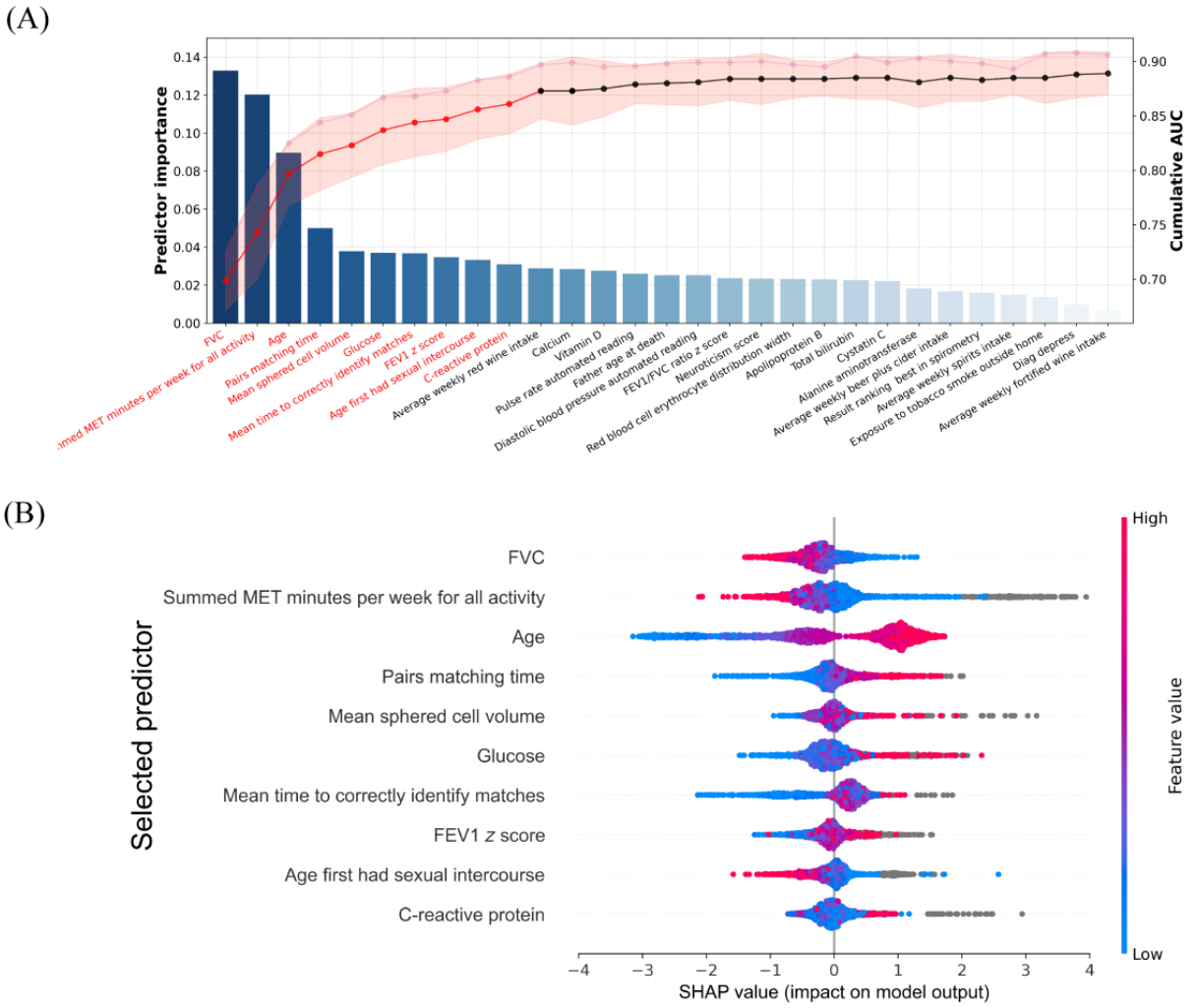
[e]PAD: peripheral artery disease.

[f]AAA: abdominal aortic aneurysm.

Data-Driven Predictors Selection

Among the 316 candidate variables, we initially selected the top 50 variables based on the LightGBM classifier and performed the hierarchical clustering to eliminate the multicollinearity [21]. As shown in the bar chart of Figure 2A, a total of 29 variables were sorted according to their importance in the prediction task. The SFS strategy was used to strike a balance between model performance (AUC on the right axis) and the number of variables selected, as depicted in the line chart. The line chart showed that the model's performance experienced a sharp increase when incorporating the first few variables and eventually reached a plateau with the inclusion of additional variables. Ultimately, the top 10 variables were chosen as the final predictors for ML model development. Their summary statistics are displayed in Table 2.

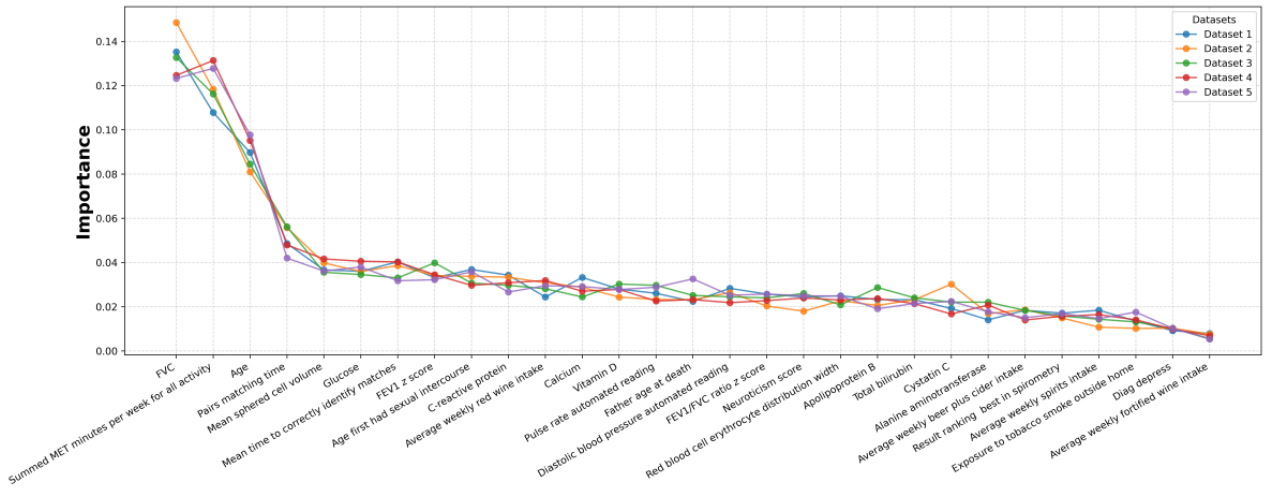
Figure 2. Predictive variable selection and interpretation on all incident dementia. (A) Sequential forward selection from a preselected predictor pool. A bar chart ranked predictor importance by their contribution to classification, while a line chart depicted cumulative AUCs with each iterative predictor inclusion. The top 10 predictors, marked red, were selected for machine learning model construction. (B) SHAP-based visualization of salient predictors. Horizontal bar widths correspond to predictor impact on model predictions, with wider ranges indicating greater influence. Predictor intensity was color-coded, graduating from blue (low) to red (high), as per the color bar on the right. The x-axis orientation signified the probability of either dementia (right) or health (left). AUC: area under the receiver operating characteristic curve; FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; MET: metabolic equivalent; SHAP: Shapley Additive Explanations.



Model Interpretation of Selected Predictors

To interpret the influence of each selected predictor, we used SHAP values and visualized them in Figure 2B. The predictors were interpreted based on value magnitude (coded by gradient colors) and tendency on the horizontal axis (indicting the likelihood of developing dementia). Take the predictor forced vital capacity (FVC) as an example. Patients with ASCVD with lower FVC values (colored blue) were more likely to develop dementia (right side) compared to those with higher FVC values (colored red). Similarly, for the remaining predictors, patients with ASCVD who spend less time engaging in activities, being older, who take longer time to complete pair matching tasks, and who have higher mean sphered cell volume (MSCV), forced expiratory volume in 1 second *z* score, C-reactive protein, and glucose levels tend to have an increased risk of developing dementia. Interestingly, we found that patients who engaged in sexual intercourse at an earlier age were more likely to develop dementia. Moreover, a 5-fold cross-validation stability analysis of feature importance was conducted. The results indicated that most features exhibited high stability across different data subsets. Figure 3 shows the distribution of key feature importance across all folds.

Figure 3. Stability analysis of feature importance across different data subsets in the model of all incident dementia. The stability of feature importance in the model of all incident dementia was assessed across different data subsets, providing insights into the robustness of identified predictors. Each line represented the variability of feature importance in a specific subset, illustrating how consistent the predictive factors are across varying conditions. FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; MET: metabolic equivalent.



Model Performance Across Different Populations and Algorithms

Compared with other ML algorithms, it can be seen in Table 3 that LightGBM demonstrated superior performance across various metrics. We used the AUC metric to evaluate the discrimination performance of the UK Biobank ASCVD risk prediction model. As depicted in Table 4, the model achieved a mean AUC of 0.866 (SD 0.027) for all incident dementia cases. Furthermore, the model demonstrated promising results for the prediction of 10-year and 5-year incident dementia, with mean AUCs of 0.876 (SD 0.024) and 0.903 (SD 0.076), respectively. The model for all incident dementia exhibited a mean accuracy of 0.883 (SD 0.01), mean sensitivity of 0.637 (SD 0.084), mean specificity of 0.914 (SD 0.012), mean precision of 0.479 (SD 0.031), and mean F_1-score of 0.546 (SD 0.043). Apart from the 5-year AD and VD predictions, the model also displayed valuable discrimination abilities for different AD and VD population groups. Specifically, the mean AUCs for all and 10-year incident AD were 0.836 (SD 0.043) and 0.828 (SD 0.112), respectively, while the mean AUCs for all and 10-year incident VD achieved 0.870 (SD 0.029) and 0.881 (SD 0.031), respectively. For specific metrics of all, 10-year, and 5-year dementia, AD, and VD predictions, please refer to Table 4 and Figures S3-S11 and S15 in Multimedia Appendix 1. We also compared the performance of the traditional Cox proportional hazards model and the LightGBM-based model in risk prediction. The AUC of all incident dementia, AD, and VD for the Cox model were 0.67, 0.67, and 0.71, respectively (Figure S1 in Multimedia Appendix 1). After competing risk analysis with death, the prediction power of the Cox model did not show a significant difference in predicting AD and VD (Tables S5-S7 and Figure S2 in Multimedia Appendix 1).

Table 3. Model performance metrics for different machine learning classifiers on all incident dementiaa.

	Accuracy, mean (SD)	Sensitivity, mean (SD)	Specificity, mean (SD)	Precision, mean (SD)	F_1-score, mean (SD)	AUC[b], mean (SD)
LightGBM[c]	0.883 (0.010)	0.637 (0.084)	0.914 (0.012)	0.479 (0.031)	0.546 (0.043)	0.866 (0.027)
XGBoost[d]	0.814 (0.01)	0.709 (0.045)	0.826 (0.01)	0.323 (0.018)	0.444 (0.024)	0.853 (0.02)
Random forest	0.829 (0.01)	0.703 (0.047)	0.844 (0.012)	0.345 (0.02)	0.463 (0.024)	0.859 (0.02)
KNN[e]	0.829 (0.014)	0.561 (0.059)	0.86 (0.012)	0.32 (0.033)	0.407 (0.041)	0.765 (0.024)
Logistic regression	0.783 (0.014)	0.627 (0.071)	0.802 (0.016)	0.27 (0.022)	0.378 (0.031)	0.795 (0.035)
ANN[f]	0.836 (0.1)	0.638 (0.145)	0.859 (0.125)	0.377 (0.174)	0.46 (0.102)	0.833 (0.025)

[a]The cutoff for binarization was established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

[c]LightGBM: light gradient boosting machine.

[d]XGBoost: extreme gradient boosting machine.

[e]KNN: k-nearest neighbor.

[f]ANN: artificial neural network.

Table 4. Model performance metrics for the prediction on different types of dementia[a].

	Accuracy, mean (SD)	Sensitivity, mean (SD)	Specificity, mean (SD)	Precision, mean (SD)	F_1-score, mean (SD)	AUC[b], mean (SD)
All incident dementia	0.883 (0.01)	0.637 (0.084)	0.914 (0.012)	0.479 (0.031)	0.546 (0.043)	0.866 (0.027)
All incident AD[c]	0.847 (0.029)	0.656 (0.125)	0.855 (0.035)	0.148 (0.01)	0.241 (0.012)	0.836 (0.043)
All incident VD[d]	0.859 (0.039)	0.683 (0.045)	0.868 (0.043)	0.206 (0.045)	0.315 (0.053)	0.870 (0.029)
10-Year incident dementia	0.877 (0.033)	0.709 (0.059)	0.886 (0.037)	0.255 (0.057)	0.374 (0.063)	0.876 (0.024)
10-Year incident AD	0.668 (0.061)	0.814 (0.214)	0.666 (0.065)	0.036 (0.006)	0.07 (0.012)	0.828 (0.112)
10-Year incident VD	0.839 (0.051)	0.757 (0.1)	0.841 (0.055)	0.098 (0.02)	0.173 (0.029)	0.881 (0.031)
5-Year incident dementia	0.939 (0.031)	0.694 (0.155)	0.942 (0.033)	0.125 (0.045)	0.211 (0.069)	0.903 (0.076)
5-Year incident AD	0.979 (0.006)	0.4 (0.5)	0.98 (0.006)	0.039 (0.049)	0.061 (0.047)	0.775 (0.243)
5-Year incident VD	0.952 (0.012)	0.471 (0.288)	0.957 (0.012)	0.106 (0.063)	0.172 (0.102)	0.803 (0.11)

[a]Cutoffs were established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

DCA demonstrated that our prediction model exhibited a higher net benefit within the threshold probability range of 0.04 to 0.97 across different time periods, significantly outperforming both the “treat all” and “treat none” baseline strategies (Figure 4). The Kolmogorov-Smirnov test was conducted to assess the calibration of the model. A *P* value greater than .05 indicates sufficient goodness of fit. Satisfactory calibrations for the development of all population groups, including 5-year or 10-year or all incident dementia, AD, and VD, were observed (Table 5 and Figures S12-S14 in Multimedia Appendix 1).

Figure 4. Clinical applicability of dementia risk prediction with a decision curve analysis. (A-C) The decision curve analysis of the UK Biobank atherosclerotic cardiovascular disease risk prediction model on all incident and 10-year incident and 5-year incident times.

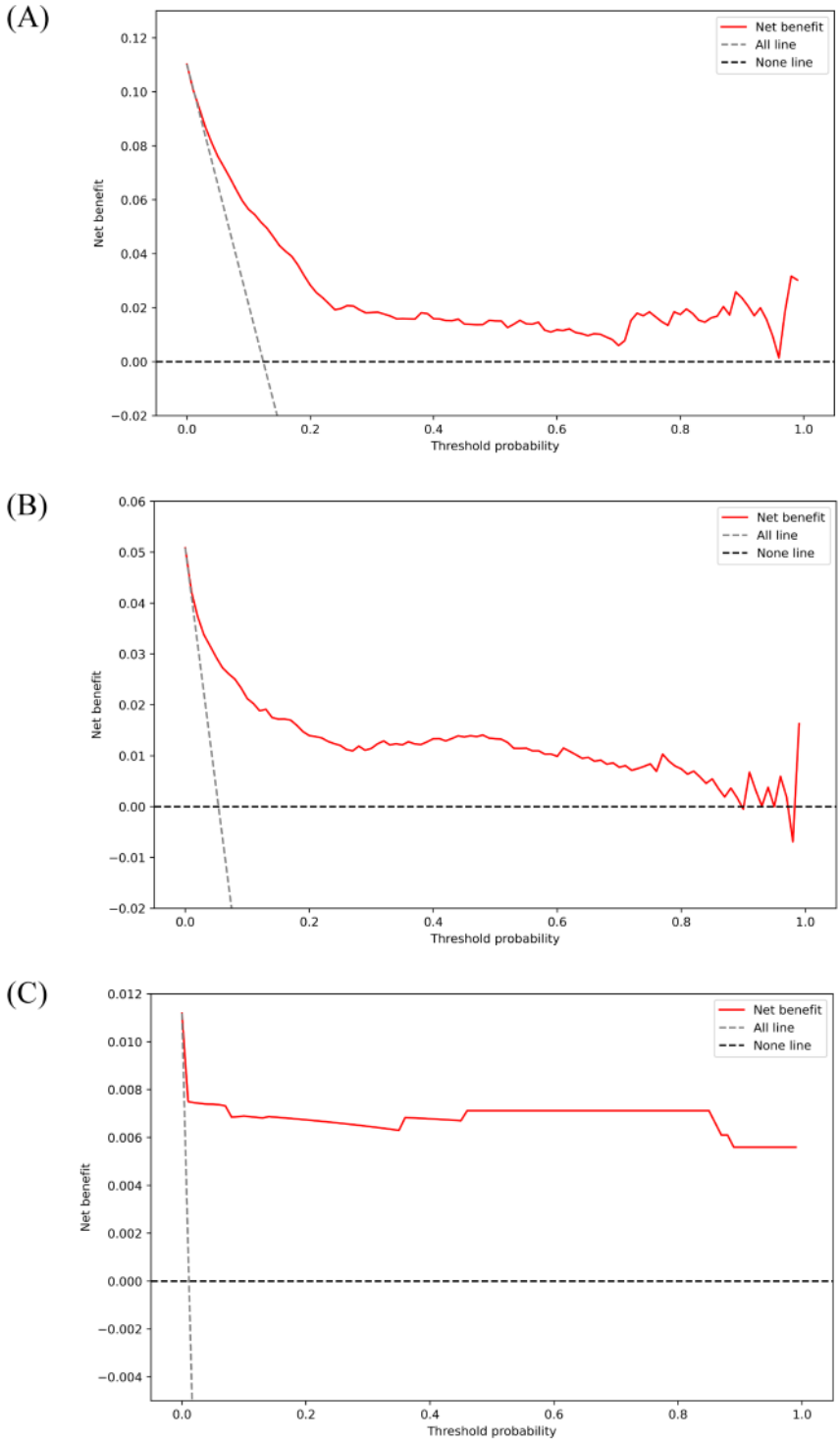


Table 5. Calibration data of the UK Biobank atherosclerotic cardiovascular disease risk prediction model[a].

Decile groups (10% quantile each)	All incident dementia[b]		All incident AD[b,c]		All incident VD[b,d]	
	Observed probability (‰)	Preserved probability (‰)	Observed probability (‰)	Preserved probability (‰)	Observed probability (‰)	Preserved probability (‰)
1	4.71	3.53	2.83	2.10	1.79	1.45
2	9.40	11.72	2.17	5.46	0.98	4.60
3	24.78	21.98	15.58	8.93	8.03	8.37
4	35.91	33.55	12.88	12.13	10.98	12.72
5	41.65	45.96	14.84	15.23	17.91	17.39
6	53.92	59.90	18.47	18.97	22.79	22.36
7	77.65	77.23	19.95	23.72	30.75	28.81
8	96.51	101.12	28.01	30.43	35.79	37.58
9	132.95	141.49	42.09	43.31	47.42	52.26
10	627.36	606.99	213.22	209.58	293.58	283.14

[a]Calibration data of the UK Biobank atherosclerotic cardiovascular disease risk prediction model on different types of dementia at all incident times. The 5-fold cross-validation strategy was performed to calculate the results. A *P* value less than .05 indicated the statistical significance of the results.

[b]Goodness-of-fit *P* value >.99.

[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

A Temporal Validation of the Constructed Model

To assess the stability and generalizability of our dementia prediction model, we used a time validation approach using data from the UK Biobank. The model was trained and validated on samples recruited between 2006 and 2009 and tested on the 2010 cohort. In the training and validation set, the model achieved a mean AUC of 0.866 (SD 0.027) in predicting all incident dementia, indicating strong discriminatory ability. When applied to the test set from 2010, the model maintained robust performance with an AUC of 0.819, suggesting good generalization to future data. Other performance metrics in the test set: accuracy was 0.851, sensitivity was 0.691, specificity was 0.866, precision was 0.315, and the F_1-score was 0.433 (Table 6).

Table 6. Model performance metrics for the prediction on the test data divided by time period[a].

	Accuracy	Sensitivity	Specificity	Precision	F_1-score	AUC[b]
All incident dementia	0.851	0.691	0.866	0.315	0.433	0.819
All incident AD[c]	0.921	0.500	0.933	0.182	0.267	0.718
All incident VD[d]	0.916	0.619	0.926	0.218	0.322	0.838
10-Year incident dementia	0.925	0.576	0.943	0.350	0.435	0.824
10-Year incident AD	0.916	0.517	0.922	0.093	0.157	0.804
10-Year incident VD	0.946	0.590	0.953	0.207	0.307	0.834
5-Year incident dementia	0.984	0.480	0.991	0.4	0.436	0.882
5-Year incident AD	0.181	1	0.181	0.001	0.002	0.605
5-Year incident VD	0.002	1	0	0.002	0.004	0.595

[a]Cutoffs were established by maximizing the Youden index (YI=sensitivity+specificity–1).

[b]AUC: area under the receiver operating characteristic curve.

[c]AD: Alzheimer disease.

[d]VD: vascular dementia.

Discussion

Principal Findings and Comparisons With Prior Work

In this study, we developed a predictive model using the LightGBM algorithm and leveraging big data from the UK Biobank to assess the risk of dementia in patients with ASCVD. To the best of our knowledge, this is the first model that uses big data to predict the risk of dementia specifically in patients with ASCVD. Our model incorporates 10 clinical predictive factors, selected based on their importance, to accurately estimate the risk of dementia. Notably, our model demonstrates particularly strong performance in predicting all-cause dementia and VD, with AUC values exceeding 0.85. Furthermore, the model effectively calibrates the predicted probabilities and aligns well with the observed event ratios, indicating its reliability and accuracy in estimating the risk of dementia in patients with ASCVD.

The ASCVD has long been recognized as one of the most significant risk factors for dementia, especially VD [22]. Although previous studies have primarily concentrated on the influence of atherosclerotic CeVDs on dementia, emerging research suggests that systemic atherosclerotic diseases beyond CeVDs also significantly contribute to the development of dementia [23]. To the best of our knowledge, no prior studies have used big data to predict the risk of dementia in patients with ASCVD. Unlike models based on variables obtained from intricate neuroimaging and neuropsychological tests, the predictors in this model are more accessible and can be applied in various clinical settings and medical institutions.

Recent studies have indicated the importance of vascular risk factors in the development of dementia, which should be taken into consideration by clinical practitioners. During the establishment of a dementia risk model in the population with ASCVD, we identified several key and distinctive risk factors that differ from other longitudinal studies. Age was identified as one of the most significant influencing factors in this study. Other factors include low exercise time, high fasting blood glucose, and reaction time including pair matching time and mean time to correctly identify matches. These risk factors have been shown to have a close relationship with the overall health of the vascular system and are critical in the development of ASCVD events [24,25]. Besides, high plasma levels of C-reactive protein at baseline were associated with a high risk of all incident dementia in this study, which is corresponded with the result of the latest research [26]. Furthermore, we found that MSCV should be considered as a new and significant factor in assessing the risk of dementia in patients with ASCVD. MSCV is primarily a parameter in hematology used to assess changes in the volume of spherocytes. Currently, there is no evidence to suggest a direct relationship between MSCV and dementia. However, overall blood health can indirectly affect cognitive function, especially in the presence of chronic anemia or other systemic diseases [27]. If an abnormal MSCV is observed in clinical practice or research, it is important to consider the patient's overall health status comprehensively, including but not limited to neurological function, to fully evaluate the patient's dementia risk. Recent studies have also gradually found that lung function may play a role in the onset of dementia by influencing brain structure [28]. Our model indicated FVC and forced expiratory volume in 1 second *z* score as protective factors in decreasing dementia risk. Finally, age first had sexual intercourse, also known as age at first intercourse (AFS), was first identified as one of the significant risk factors for dementia in populations with ASCVD, which is interesting and worth attention. Compared to other biological traits, reproductive behaviors, especially sexual factors, have long been neglected when it comes to the study of CVDs and neurological disorders. Recent studies show that the earlier the age of first sexual intercourse, the higher the likelihood of developing hypertension and CVDs [29,30]. The specific processes driving this relationship are not yet fully understood, but they may include a mix of environmental and genetic influences. For instance, early sexual activities are often accompanied by adverse environmental factors, including lower educational attainment, increased smoking and alcohol consumption, and the use of illicit drugs [31], which are all closely associated with CVDs. Recent studies have identified a causal relationship between AFS and CVDs at the genetic level [30,32]. Although AFS has no direct impact on dementia, it might be induced from our study that AFS is indirectly related to dementia through intriguing CVDs especially in patients with ASCVD.

In our study, we performed a stability analysis of feature importance using 5-fold cross-validation to ensure the robustness of the identified predictors. The analysis revealed that several key features consistently ranked highly across all folds, indicating their strong and reliable association with the outcome. Specifically, FVC, summed metabolic equivalent minutes per week, age, and pair matching time maintained relatively high importance scores in every subset, underscoring their role as robust predictors. Although these 4 variables exhibited significant variation compared to other variables, the impact of this variation is minimal relative to their importance.

We acknowledge that the Cox proportional hazards model, as a mature and interpretable method, holds a significant position in survival analysis. However, our study results show that the LightGBM-based ML model significantly outperforms the Cox model in predictive performance metrics such as AUC, demonstrating its advantages in handling nonlinear relationships and complex interactions between variables. LightGBM effectively captures patterns in high-dimensional data, thereby enhancing the accuracy of risk prediction. Although ML models face certain challenges in terms of computational resources and interpretability, their substantial improvement in predictive performance illustrates their added value in practical applications. Future research could explore combining traditional Cox models with ML methods to balance predictive performance and model interpretability, thereby meeting diverse clinical application needs. Besides, our study indicated that mortality had a minimal impact on the primary outcomes, and the overall conclusions of the study remained largely unchanged. This suggested that the prediction power of our model remained robust even after accounting for competing risks.

The time validation results demonstrate that our dementia prediction model maintains robust performance across different

temporal cohorts. Specifically, in the population of all incident dementia, the AUC remained consistently high, with a slight decrease from 0.866 in the training set to 0.819 in the test set. Similarly, other performance metrics such as accuracy, precision, recall, and F_1-score showed only minor declines over time. This stability suggests that the model effectively captures underlying patterns associated with dementia risk that are persistent across the studied time periods. The consistency of performance metrics across the training and test sets indicates that the model's predictive capabilities are not significantly affected by temporal shifts in the data. The stable performance of the model over different time periods enhances its long-term applicability in clinical and public health settings. A model that maintains its predictive accuracy over time is invaluable for ongoing and future dementia screening programs, enabling early identification of at-risk individuals with confidence in its sustained reliability. However, it is also essential to acknowledge the significant performance decline observed in the test set when predicting the risk of 5-year AD and VD, which may be attributed to the low prevalence in patients with ASCVD during a relative short period. To ensure continued efficacy, periodic retraining and validation of the model with new data may be necessary. This approach would allow the model to adapt to any emerging trends or shifts in risk factors that may influence dementia incidence over time.

In clinical practice, the LightGBM model can be applied during the initial diagnosis or follow-up stages to early identify individuals with high dementia risk among patients with ASCVD, thus promoting timely intervention and treatment. For instance, using this predictive model during a patient's initial visit can assist physicians in swiftly identifying patients at high risk and arranging further diagnostic tests or interventions. Moreover, integrating the model's predictions into electronic health record systems can generate alerts and recommend further evaluations, thereby enhancing diagnostic accuracy and personalized treatment plans. To effectively communicate the predicted risk, doctors should use easily understandable language to explain the model results and their implications to the patient while also providing clear next steps and support resources to alleviate patient anxiety. Consider a hypothetical example: a 55-year-old male patient with hypertension and high cholesterol who recently experienced a heart attack. After using the LightGBM model for assessment, the results indicate a high dementia risk. Based on this, the doctor decides to schedule detailed cognitive function tests and recommends a comprehensive plan that includes cognitive training, a healthy diet, and regular exercise. Through these interventions, the patient can better manage his cardiovascular health while taking steps to reduce the likelihood of developing dementia.

However, the application of such predictive models raises potential ethical issues. First, there may be prediction bias due to training data, leading to unequal care, so continuous monitoring and validation are needed to ensure fairness. Second, doctors need to carefully communicate the model's predictions to avoid causing unnecessary anxiety for patients. Additionally, patient data use should require clear consent and ensure privacy protection. Finally, caution against overreliance on model predictions is necessary, with doctors maintaining primary responsibility for care decisions.

Limitations

Several limitations should be considered when interpreting the results. First, our study focused on a specific population of patients with ASCVD. Due to the relatively small sample size, we observed a lower AUC value when predicting the 5-year incidence rates of dementia, especially for AD and VD incidence over a 5-year period. This issue can be addressed by further expanding the sample size. Additionally, this study primarily used samples of European descent, which may restrict the generalizability of our findings to other populations. Genetic, environmental, and lifestyle differences across diverse ethnic groups could influence the model's performance and predictive accuracy. The limited diversity of the sample may affect the model's applicability to non-European populations. To ensure broader relevance and robustness, future research should include diverse ethnic backgrounds to validate and potentially refine the model for varied demographic groups. While the time validation results are promising, the model currently relies on static features collected at baseline. Integrating longitudinal data and time-varying covariates could potentially improve predictive performance and adaptability over extended periods. Despite incorporating death as a primary competing risk, there might still be other unrecognized or unadjusted competing factors, such as other chronic diseases or lifestyle changes, that could influence the results to some extent. The application of competing risk models relied on the correct specification of models and assumptions; any biases in model setup might affect the accuracy of the analysis. Therefore, future research should further explore additional potential competing risk factors and use more sophisticated statistical methods to comprehensively assess the prediction power. Regarding the ML algorithms chosen for our study, the use of LightGBM may lead to data overfitting because it generates deep decision trees. To mitigate overfitting, a maximum depth limit should be imposed during the use of LightGBM. Furthermore, it is important to acknowledge that LightGBM is a bias-based algorithm and can be sensitive to noise in data processing, which may potentially affect the final data analysis results. Additionally, it should be noted that the predictive variables identified in this study were derived from data-driven analytical models, which may induce some bias compared to actual clinical diagnostic and treatment experiences. While advanced predictive models and results have been obtained, their applicability to clinical practice remains uncertain. Therefore, future research should focus on validating the analysis results using other independent cohorts with larger sample sizes and extending the study methodology to populations from different countries, regions, and ethnicities. The integration of clinical practice experiences will contribute to the development of more universally applicable and practical models.

Conclusions

This study has identified several practical and novel predictors for dementia screening in patients with ASCVD. It is worthy of testing and evaluating the applicability of these factors in clinical practice. Future studies should focus on investigating

whether intervening in these factors can help prevent the incidence of dementia in patients with ASCVD. By exploring these possibilities, we can potentially improve the management and outcomes of patients with ASCVD and reduce the burden of dementia in this population.

Acknowledgments

This study was funded by grants from the Natural Science Foundation of Shanghai (23JS1400500 and 22ZR1420500), the Shanghai ShenKang Hospital Development Center (SHDC2020CR1038B), and 234 discipline peak climbing program of Shanghai Changhai Hospital (2019YXK013). The funding sources were not involved in the study's design and execution; data collection, management, analysis, and interpretation; manuscript drafting, review, or approval; nor the decision to submit the manuscript for publication. This study was conducted by the use of the UK Biobank Resource (application 97366). The authors appreciate all the participants and researchers from the UK Biobank.

Data Availability

All data used during this study are available from the UK Biobank Resource. As the datasets are the property of UK Biobank, they are not available for direct request. Researchers interested in the data can apply for access through the UK Biobank Access Management System.

Authors' Contributions

XS, XB, and BD designed the project and were responsible for the integrity of data and accuracy of analysis. HM, SL, and YL played a major role in the model's development and validation as well as statistical analysis. ZG, BD, XJ, and XG contributed to the discussion of results and drafted the manuscript. All authors contributed to the paper and took responsibility for submitting it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
UK Biobank atherosclerotic cardiovascular disease study supplementary tables, figures, and methodological details.
[DOCX File , 5010 KB - aging_v8i1e64148_app1.docx]

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020;76(25):2982-3021 [FREE Full text] [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1736-1788 [FREE Full text] [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
3. Zhao D, Liu J, Wang M, Zhang X, Zhou M. Epidemiology of cardiovascular disease in China: current features and implications. Nat Rev Cardiol 2019;16(4):203-212. [doi: 10.1038/s41569-018-0119-4] [Medline: 30467329]
4. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
5. Winstein CJ, Stein J, Arena R, Bates B, Cherney LR, Cramer SC, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
6. Sabia S, Dugravot A, Dartigues JF, Abell J, Elbaz A, Kivimäki M, et al. Physical activity, cognitive decline, and risk of dementia: 28 year follow-up of Whitehall II cohort study. BMJ 2017;357:j2709 [FREE Full text] [doi: 10.1136/bmj.j2709] [Medline: 28642251]
7. Sabia S, Fayosse A, Dumurgier J, Schnitzler A, Empana J, Ebmeier KP, et al. Association of ideal cardiovascular health at age 50 with incidence of dementia: 25 year follow-up of Whitehall II cohort study. BMJ 2019;366:l4414 [FREE Full text] [doi: 10.1136/bmj.l4414] [Medline: 31391187]
8. Schwarzinger M, Pollock BG, Hasan OSM, Dufouil C, Rehm J, QalyDays Study Group. Contribution of alcohol use disorders to the burden of dementia in France 2008-13: a nationwide retrospective cohort study. Lancet Public Health 2018;3(3):e124-e132 [FREE Full text] [doi: 10.1016/S2468-2667(18)30022-7] [Medline: 29475810]

9. Veronese N, Facchini S, Stubbs B, Luchini C, Solmi M, Manzato E, et al. Weight loss is associated with improvements in cognitive function among overweight and obese people: a systematic review and meta-analysis. Neurosci Biobehav Rev 2017;72:87-94 [FREE Full text] [doi: 10.1016/j.neubiorev.2016.11.017] [Medline: 27890688]
10. Nordestgaard LT, Christoffersen M, Frikke-Schmidt R. Shared risk factors between dementia and atherosclerotic cardiovascular disease. Int J Mol Sci 2022;23(17):9777 [FREE Full text] [doi: 10.3390/ijms23179777] [Medline: 36077172]
11. Deo RC. Machine learning in medicine. Circulation 2015;132(20):1920-1930 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
12. James C, Ranson JM, Everson R, Llewellyn DJ. Performance of machine learning algorithms for predicting progression to dementia in memory clinic patients. JAMA Netw Open 2021;4(12):e2136553 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.36553] [Medline: 34913981]
13. Zhan Y, Chen K, Wu X, Zhang D, Zhang J, Yao L, et al. Identification of conversion from normal elderly cognition to Alzheimer's disease using multimodal support vector machine. J Alzheimers Dis 2015;47(4):1057-1067 [FREE Full text] [doi: 10.3233/JAD-142820] [Medline: 26401783]
14. Park JH, Cho HE, Kim JH, Wall MM, Stern Y, Lim H, et al. Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. NPJ Digit Med 2020;3:46 [FREE Full text] [doi: 10.1038/s41746-020-0256-0] [Medline: 32258428]
15. Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, et al. Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLoS One 2011;6(7):e21896 [FREE Full text] [doi: 10.1371/journal.pone.0021896] [Medline: 21814561]
16. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12(3):e1001779 [FREE Full text] [doi: 10.1371/journal.pmed.1001779] [Medline: 25826379]
17. Hageman SHJ, McKay AJ, Ueda P, Gunn LH, Jernberg T, Hagström E, et al. Estimation of recurrent atherosclerotic cardiovascular event risk in patients with established cardiovascular disease: the updated SMART2 algorithm. Eur Heart J 2022;43(18):1715-1727 [FREE Full text] [doi: 10.1093/eurheartj/ehac056] [Medline: 35165703]
18. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;350:g7594 [FREE Full text] [doi: 10.1136/bmj.g7594] [Medline: 25569120]
19. Wang J, Qiu J, Zhu T, Zeng Y, Yang H, Shang Y, et al. Prediction of suicidal behaviors in the middle-aged population: machine learning analyses of UK Biobank. JMIR Public Health Surveill 2023;9:e43419 [FREE Full text] [doi: 10.2196/43419] [Medline: 36805366]
20. Li Q, Yang X, Xu J, Guo Y, He X, Hu H, et al. Early prediction of Alzheimer's disease and related dementias using real-world electronic health records. Alzheimers Dement 2023;19(8):3506-3518 [FREE Full text] [doi: 10.1002/alz.12967] [Medline: 36815661]
21. Martin SA, Townend FJ, Barkhof F, Cole JH. Interpretable machine learning for dementia: a systematic review. Alzheimers Dement 2023;19(5):2135-2149 [FREE Full text] [doi: 10.1002/alz.12948] [Medline: 36735865]
22. Iadecola C. Revisiting atherosclerosis and dementia. Nat Neurosci 2020;23(6):691-692 [FREE Full text] [doi: 10.1038/s41593-020-0626-6] [Medline: 32424283]
23. Zlokovic BV, Gottesman RF, Bernstein KE, Seshadri S, McKee A, Snyder H, et al. Vascular contributions to cognitive impairment and dementia (VCID): a report from the 2018 National Heart, Lung, and Blood Institute and National Institute of Neurological Disorders and Stroke Workshop. Alzheimers Dement 2020;16(12):1714-1733. [doi: 10.1002/alz.12157] [Medline: 33030307]
24. Meyer-Lindemann U, Moggio A, Dutsch A, Kessler T, Sager HB. The impact of exercise on immunity, metabolism, and atherosclerosis. Int J Mol Sci 2023;24(4):3394 [FREE Full text] [doi: 10.3390/ijms24043394] [Medline: 36834808]
25. Stone NJ, Smith SC, Orringer CE, Rigotti NA, Navar AM, Khan SS, et al. Managing atherosclerotic cardiovascular risk in young adults: JACC state-of-the-art review. J Am Coll Cardiol 2022;79(8):819-836 [FREE Full text] [doi: 10.1016/j.jacc.2021.12.016] [Medline: 35210038]
26. Tachibana A, Iga JI, Ozaki T, Yoshida T, Yoshino Y, Shimizu H, et al. Serum high-sensitivity C-reactive protein and dementia in a community-dwelling Japanese older population (JPSC-AD). Sci Rep 2024;14(1):7374 [FREE Full text] [doi: 10.1038/s41598-024-57922-1] [Medline: 38548879]
27. Qiang YX, Deng YT, Zhang YR, Wang H, Zhang W, Dong Q, et al. Associations of blood cell indices and anemia with risk of incident dementia: a prospective cohort study of 313,448 participants. Alzheimers Dement 2023;19(9):3965-3976. [doi: 10.1002/alz.13088] [Medline: 37102212]
28. Ma YH, Shen LX, Li YZ, Leng Y, Yang L, Chen S, et al. Lung function and risk of incident dementia: a prospective cohort study of 431,834 individuals. Brain Behav Immun 2023;109:321-330. [doi: 10.1016/j.bbi.2023.02.009] [Medline: 36796705]
29. Ngueta G, Ndjaboue R. Early sexual experience and hypertension in US adults: results from the National Health and Nutrition Examination Survey 2001-2016. J Hypertens 2018;36(12):2414-2419. [doi: 10.1097/HJH.0000000000001821] [Medline: 29957720]

30. Zhuo C, Chen L, Wang Q, Cai H, Lin Z, Pan H, et al. Association of age at first sexual intercourse and lifetime number of sexual partners with cardiovascular diseases: a bi-directional Mendelian randomization study. Front Cardiovasc Med 2023;10:1267906 [FREE Full text] [doi: 10.3389/fcvm.2023.1267906] [Medline: 38146444]
31. Boisvert I, Boislard MA, Poulin F. Early sexual onset and alcohol use and misuse from adolescence into young adulthood. J Adolesc Health 2017;61(4):514-520. [doi: 10.1016/j.jadohealth.2017.04.013] [Medline: 28732719]
32. Mills MC, Tropf FC, Brazel DM, van Zuydam N, Vaez A, Pers TH, et al. Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. Nat Hum Behav 2021;5(12):1717-1730 [FREE Full text] [doi: 10.1038/s41562-021-01135-3] [Medline: 34211149]

Abbreviations

AD: Alzheimer disease
AFS: age at first intercourse
ASCVD: atherosclerotic cardiovascular disease
AUC: area under the receiver operating characteristic curve
CeVD: cerebrovascular disease
CVD: cardiovascular disease
DCA: decision curve analysis
FVC: forced vital capacity
LightGBM: light gradient boosting machine
ML: machine learning
MSCV: mean sphered cell volume
SFS: sequential forward selection
SHAP: Shapley Additive Explanations
VD: vascular dementia

Edited by G Seçkin; submitted 10.07.24; peer-reviewed by T Yang, L Mercado Diaz; comments to author 22.08.24; revised version received 31.10.24; accepted 20.01.25; published 26.02.25.

Please cite as:
Gu Z, Liu S, Ma H, Long Y, Jiao X, Gao X, Du B, Bi X, Shi X
Estimation of Machine Learning–Based Models to Predict Dementia Risk in Patients With Atherosclerotic Cardiovascular Diseases: UK Biobank Study
JMIR Aging 2025;8:e64148
URL: https://aging.jmir.org/2025/1/e64148
doi:10.2196/64148
PMID:

Original Paper

Development and Validation of a Brain Aging Biomarker in Middle-Aged and Older Adults: Deep Learning Approach

Zihan Li[1*], MS; Jun Li[2*], MD; Jiahui Li[1*], BM; Mengying Wang[1], MM; Andi Xu[1], MM; Yushu Huang[1], MM; Qi Yu[1], MM; Lingzhi Zhang[1], MM; Yingjun Li[3], PhD; Zilin Li[4], PhD; Xifeng Wu[1,5], MD, PhD; Jiajun Bu[6], PhD; Wenyuan Li[1,5], ScD

[1]Center for Clinical Big Data and Analytics, The Second Affiliated Hospital and Department of Big Data in Health Science, School of Public Health, Zhejiang University School of Medicine, Zhejiang University, Hangzhou, China

[2]Department of Radiology, The First Affiliated Hospital, Zhengzhou University, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

[3]Department of Epidemiology and Health Statistics, School of Public Health, Hangzhou Medical College, Hangzhou Medical College, Hangzhou, China

[4]School of Mathematics and Statistics, Northeast Normal University, Northeast Normal University, Changchun, China

[5]Zhejiang Provincial Key Laboratory of Intelligent Preventive Medicine, Hangzhou, China

[6]Zhejiang Key Laboratory of Accessible Perception and Intelligent Systems, College of Computer Science and Technology, Zhejiang University, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Wenyuan Li, ScD
Center for Clinical Big Data and Analytics, The Second Affiliated Hospital and Department of Big Data in Health Science, School of Public Health
Zhejiang University School of Medicine
Zhejiang University
Zhejiang Campus
866 Yuhangtang Road
Hangzhou, 310058
China
Phone: 86 19521623869
Email: wenyuanli@zju.edu.cn

Abstract

Background: Precise assessment of brain aging is crucial for early detection of neurodegenerative disorders and aiding clinical practice. Existing magnetic resonance imaging (MRI)–based methods excel in this task, but they still have room for improvement in capturing local morphological variations across brain regions and preserving the inherent neurobiological topological structures.

Objective: To develop and validate a deep learning framework incorporating both connectivity and complexity for accurate brain aging estimation, facilitating early identification of neurodegenerative diseases.

Methods: We used 5889 T1-weighted MRI scans from the Alzheimer's Disease Neuroimaging Initiative dataset. We proposed a novel brain vision graph neural network (BVGN), incorporating neurobiologically informed feature extraction modules and global association mechanisms to provide a sensitive deep learning–based imaging biomarker. Model performance was evaluated using mean absolute error (MAE) against benchmark models, while generalization capability was further validated on an external UK Biobank dataset. We calculated the brain age gap across distinct cognitive states and conducted multiple logistic regressions to compare its discriminative capacity against conventional cognitive-related variables in distinguishing cognitively normal (CN) and mild cognitive impairment (MCI) states. Longitudinal track, Cox regression, and Kaplan-Meier plots were used to investigate the longitudinal performance of the brain age gap.

Results: The BVGN model achieved an MAE of 2.39 years, surpassing current state-of-the-art approaches while obtaining an interpretable saliency map and graph theory supported by medical evidence. Furthermore, its performance was validated on the UK Biobank cohort (N=34,352) with an MAE of 2.49 years. The brain age gap derived from BVGN exhibited significant difference across cognitive states (CN vs MCI vs Alzheimer disease; $P<.001$), and demonstrated the highest discriminative capacity between CN and MCI than general cognitive assessments, brain volume features, and apolipoprotein E4 carriage (area under the receiver operating characteristic curve [AUC] of 0.885 vs AUC ranging from 0.646 to 0.815). Brain age gap exhibited clinical feasibility

combined with Functional Activities Questionnaire, with improved discriminative capacity in models achieving lower MAEs (AUC of 0.945 vs 0.923 and 0.911; AUC of 0.935 vs 0.900 and 0.881). An increasing brain age gap identified by BVGN may indicate underlying pathological changes in the CN to MCI progression, with each unit increase linked to a 55% (hazard ratio=1.55, 95% CI 1.13-2.13; *P*=.006) higher risk of cognitive decline in individuals who are CN and a 29% (hazard ratio=1.29, 95% CI 1.09-1.51; *P*=.002) increase in individuals with MCI.

Conclusions: BVGN offers a precise framework for brain aging assessment, demonstrates strong generalization on an external large-scale dataset, and proposes novel interpretability strategies to elucidate multiregional cooperative aging patterns. The brain age gap derived from BVGN is validated as a sensitive biomarker for early identification of MCI and predicting cognitive decline, offering substantial potential for clinical applications.

(JMIR Aging 2025;8:e73004) doi:10.2196/73004

KEYWORDS

brain aging; deep learning; magnetic resonance imaging; MRI; imaging biomarker

Introduction

Background

As human life expectancy steadily rises, the global population aged ≥60 years is projected to nearly double by 2050, reaching an estimated 2.1 billion [1]. This demographic shift presents substantial burdens to society, including rising health care costs and caregiving demands [2]. In the area of aging, dementia accounts for 11.2% of disability years, surpassing that caused by strokes (9.5%), musculoskeletal disorders (8.9%), cardiovascular diseases (5%), and all types of cancer (2.4%) [3]. Scientific investigations have illuminated the association between an aging brain and neurodegenerative conditions, such as Alzheimer disease (AD), by revealing the molecular and cellular mechanisms, particularly mitochondrial impairment [4,5]. There is an urgent demand for accurate assessment of brain aging to enable early detection of age-related neurodegenerative diseases, with neuroimaging-based brain age estimation methods gaining popularity in recent years [6].

Given its exceptional soft tissue contrast, noninvasive nature, and multimodal capabilities, magnetic resonance imaging (MRI) is the preferred method for timely visualizing brain structure and subtle lesions [7,8]. Brain age gap estimation [9], a quantified MRI-based score, has been considered a reliable brain age scale [10,11], with numerous studies associating a positive brain age gap with brain abnormalities and mortality [12,13]. The typical approach in this field involves constructing a standardized model using structural MRI data from healthy individuals. This model is then used to evaluate neuroanatomical deviations from the norm in new participants, providing insight into underlying brain abnormalities [14]. Recent publications have demonstrated the efficacy of traditional methods for brain age gap estimation, such as support vector regression and relevance vector regression, with mean absolute errors (MAEs) ranging from 2.6 years to 7.7 years [15] and 3.7 years to 4.7 years [16], respectively. However, these approaches often rely on supervised learning, with inherent limitations related to manual labeling and model performance influenced by dataset bias [14].

In recent years, the emergence and widespread adoption of deep learning (DL) techniques, which require no manual annotation and excel in handling high-dimensional data [17], have propelled model precision to unprecedented heights, achieving MAEs of <3 years [18-20]. For instance, BrainAgeNeXt [21] achieved an MAE of 2.78 years. Convolutional neural networks (CNNs) are lauded for their adeptness in automatic feature extraction from MRI images but are limited by the gradient explosion or vanishing issue in deeper layers. Residual neural networks (ResNets) have made significant strides by introducing skip connections to mitigate these training challenges, albeit at the cost of increased computational complexity. Transformer-based models represent a paradigm shift with their self-attention mechanisms adept at capturing long-range dependencies, but they demand substantial data for optimal training. Ensemble DL approaches have garnered attention for their ability to synthesize the predictions of multiple models, thereby enhancing the robustness and accuracy of brain age estimation. This prediction strategy, while beneficial for improving generalization, incurs the surcharge of increased training complexity. However, most DL-based approaches regard brain MRI as a cube constructed of voxels and then transfer models that perform well in natural images or videos directly to this task. This approach neglects the unique biological characteristics of the brain, which distinguish it from natural images [22], particularly in morphology [23]. Moreover, the human connectome exhibits intricate higher-order connectivity, encompassing spatial, functional, and sex-specific distinctions [24] among different brain voxel regions. Common CNN architectures designed for natural images lack the capacity to modify their topologies [25]. Currently, there is a notable absence of a DL framework that simultaneously considers both brain morphology and voxel regions connectivity for assessing brain aging. In addition, there is insufficient evidence to compare and quantify the clinical application value of the brain age gap and the relationship between narrowing MAEs with improving clinical application performance.

This Study

Here we proposed a novel graph-based DL framework, a brain vision graph neural network (BVGN), which uses deformable kernels to extract brain morphological features and incorporated a graph neural network (GNN) to characterize voxel regions connectivity. We pioneered the adaptation of vision GNN [26], a backbone widely researched in natural image processing, to the domain of brain aging by extending its framework to 3D MRI. The BVGN leveraged an individual's standard

T1-weighted MRI as input and generated the corresponding neuroimaging biomarker-brain age gap to evaluate brain aging. Subsequently, we constructed multiple logistics regression models, each incorporating different types of input variables, to classify data samples labeled as cognitively normal (CN) and with mild cognitive impairment (MCI). We aimed to compare and validate the effects of brain age gap against conventional variables, including demographic factors, brain volume features, apolipoprotein E4 (APOE4) carriage, and general cognitive scales, in distinguishing CN and MCI states. In addition, we assessed the combined utility of the brain age gap with the Functional Activities Questionnaire (FAQ) cognitive assessment and the impact of lower MAEs. Cox proportional hazards regressions were performed separately using data from participants classified as CN and those with MCI at baseline to evaluate the potential of brain age gap as a risk biomarker for cognitive decline over a 2-year follow-up period.

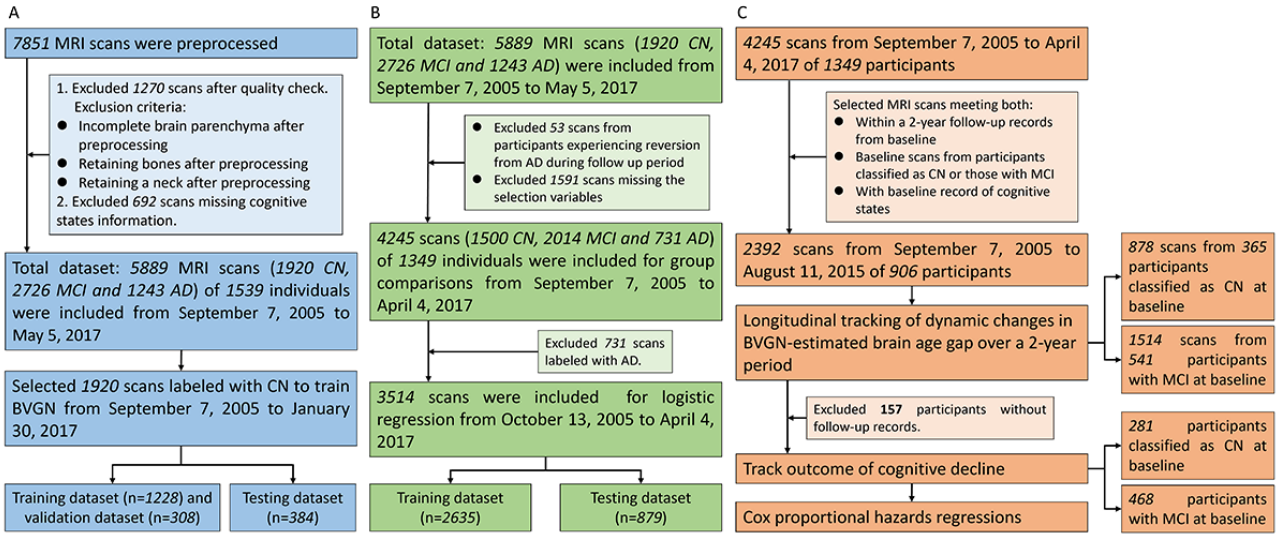
Methods

Dataset, Study Design, and Inclusion Criteria

Data used in the preparation of this paper were obtained from the database of Alzheimer's Disease Neuroimaging Initiative (ADNI) studies [27-29]. The ADNI was launched in 2003 as a public-private partnership, led by principal investigator Michael W Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

This study gathered 7851 T1-weighted MRI scans from ADNI-1, ADNI-2, and ADNI-GO (Grand Opportunities) studies. The research design comprised 2 sequential phases: development and validation of the BVGN model for accurate brain age estimation, and clinical effect validation of the BVGN-estimated brain age gap metric by integrated cross-sectional and longitudinal analysis. For BVGN model development, we specifically chose 5889 scans from 1539 individuals that contained cognitive states labels and met both the ADNI MRI scanner protocols [30,31] and our strict training criteria (ie, presenting complete structural information after preprocessing). Data regarding the patient's cognitive states were retrieved from the ADNI web portal [32]. Only 1920 MRI scans of participants classified as CN were used for training the BVGN model (Figure 1A). To ensure comprehensive model development and evaluation, we divided the dataset consisting of participants classified as CN into training set with validation set and test set in a ratio of 8:2. The training set and validation set were used for model training and hyperparameter setting, while the test set was used to validate the performance of the model. A subset of the UK Biobank (UKB) dataset [33] was used for external validation to evaluate model generalizability. We curated 34,352 MRI scans, partitioned into training set (n=21,985, 64%), validation set (n=5497, 16%), and test set (n=6870, 20%).

Figure 1. Flow diagram to show study design and inclusion criteria. (A) Brain vision graph neural network (BVGN) training and testing part; (B) cross-sectional validation part; (C) longitudinal validation part. AD: Alzheimer disease; CN: cognitively normal; MCI: mild cognitive impairment; MRI: magnetic resonance imaging.



Given the progressive nature of dementia caused by neurodegenerative disorders, we excluded scans from participants who experienced reversion from AD during the follow-up period (Figure 1B). Following feature selection, 4245 MRI scans from 1349 participants were retained for cross-sectional analysis. All scans were used for intergroup comparisons, while only those from participants classified as CN and those with MCI were used for logistic regression analysis aiming at binary classification between CN and MCI. To validate the performance of logistic regression models, the dataset was split into training and test sets at an 8:2 ratio.

We combined various MRI scans from the same individual for longitudinal analysis (Figure 1C). Specifically, we used a dual-component longitudinal design to investigate the trajectories of BVGN-estimated brain age gap across cognitive progression groups and quantify cognitive decline risks through Cox proportional hazard models. For the brain age gap trajectory analysis, cognitive progression categories (CN-CN, CN-MCI or CN-AD, MCI-CN or MCI-MCI, and MCI-AD) were stratified

based on recorded cognitive states from baseline to a 2-year follow-up. The inclusion criteria required participants to be classified as either CN or having MCI at the baseline, yielding 1514 scans from 541 participants with MCI and 878 scans from 365 participants classified as CN.

To evaluate time-dependent cognitive decline, we conducted Cox proportional hazards regression stratified by baseline groups. We excluded 157 participants without follow-up data and performed separate analyses for the baseline CN group (tracking MCI conversion) and MCI group (tracking AD conversion). The final analysis included 749 participants, comprising 281 (37.5%) individuals classified as CN and 468 (62.5%) individuals with MCI, with the time-to-event defined as the interval between baseline MRI and first observed cognitive transition.

Quality Control and Data Preprocessing

For the ADNI dataset, each series within every examination was subjected to a rigorous quality control process at the Mayo Clinic. This involved 2 distinct levels of scrutiny—compliance with protocol-specific parameters and an assessment of the series-specific quality, such as the participant's movement and the extent of anatomical coverage. The quality of the scans was evaluated by trained analysts who assigned a subjective grade—scores of 1 to 3 were considered satisfactory, while a score of 4 indicated a failure, rendering the scans unusable. While the UKB dataset was only used to test the model's generalizability, all data also underwent standardized quality control procedures [34].

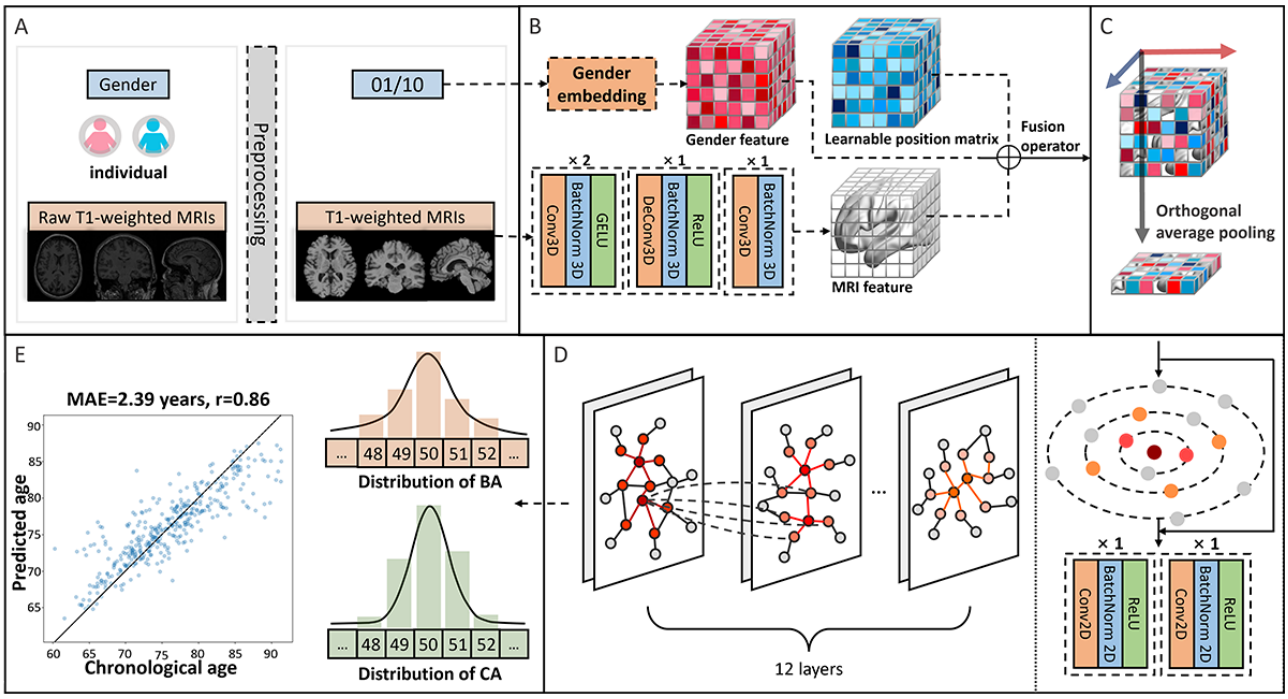
The acquisition protocols of different datasets required data preprocessing to ensure compatibility. The entire MRI preprocessing pipeline could be divided into 4 steps. First, the raw MRI data in neuroimaging informatics technology initiative format underwent minimum-maximum normalization. Second, nonbrain tissue, such as the skull and neck, was removed. In the third step, each MRI was registered from its native space to the standard Montreal Neurological Institute 152 1-mm^3 template [35] chosen for registration in our study. Finally, central cropping was performed on the complete MRI volume, resulting in a final output size of 160×196×160 mm^3. In addition, further preprocessing steps were conducted to exclude MRIs with incomplete brain structures, ensuring dataset quality. The preprocessing pipeline was implemented using SimpleITK [36] and FMRIB software library [37].

During the logistic and Cox regression modeling phase, *z* score normalization was applied to numerical variables, while categorical variables were transformed using one-hot encoding.

BVGN Model Development

Owing to the need for a large number of high-quality handcrafted features in traditional methods [14] and because inductive bias to brain structures is not effectively transferable from natural images when applying classical DL architectures, there are limitations in extracting features from irregular objects, such as cortical sulcus and gyrus, as well as parenchyma using grid-like convolution kernels that perform weighted summation of pixel values at the centroid and its neighbors [25,38,39]. Furthermore, CNN architectures lack the ability to modify topological relationships between voxel regions [22]. To address these issues, we proposed an end-to-end framework called BVGN (Figure 2), which used standardized MRIs to estimate brain aging of individuals. To be more specific, BVGN makes it easier to extract irregularly shaped structures in brain MRI by superimposing deformable convolution kernels. A graph CNN was used to model the relationship between different regions.

Figure 2. Overview of the brain vision graph neural network model. (A) A preprocessing pipeline was implemented, which involved converting gender into 1-hot encoding and transforming raw magnetic resonance imaging (MRI) data into brain tissue located in a standard space. (B) Discretization layer, the voxel space of brain is divided by the discretization layer, which uses a combination of multiple convolution modules and deformable convolution modules. Subsequently, positional features and gender features are integrated into cubes. (C) Orthogonal average pooling module (coronal axis, sagittal axis, transverse axis, the 3D feature cube is transformed into a 2D feature map). (D) Stacked graph neural network modules. (E) During training, the image on the right illustrates that the supervised signal of back-propagation is computed by evaluating the Kullback-Leibler divergence between estimated brain age and chronological age; whereas the image on the left depicts the model's performance. BA: brain age; CA: chronological age.



The BVGN framework consists of 3 components: discretization layer, stacked GNN layer, and aging distribution prediction header. Specifically, standardized MRIs were discretized into multiple cubes using nonoverlapping multiple deformable convolutional modules [38], where each cube's feature vectors contained all information within the receptive field range of deformable convolutional modules. In addition, an individual's sex and position information for each cube was incorporated into the feature map by passing through a projection module, modal-type embedding [40] and initializing a learnable parameter matrix [41]. An orthogonal average pooling module was implemented to project the 3D volumetric feature tensor along 3 mutually perpendicular anatomical axes (coronal, sagittal, and transverse). This dimensionality reduction operation transforms the 3D spatial representation into a 2D feature map while preserving critical spatial information across orthogonal planes. Detailed explanations of Figure 2 are provided in Multimedia Appendix 1.

The feature map was subsequently transformed into a graph using Euclidean distance and nearest neighbors as the basis, which was then input into a stacked GNN [26,42]. Importantly, the dynamic construction of the graph during forward processing relied on assessing similarity between feature maps at different depths. Ultimately, aging probability distributions [43,44] were obtained in the prediction header by integrating the feature map through bottlenecked convolutional kernels [45], based on predefined age ranges. Inspired by vision GNN [26], we implemented 2 distinct neural network architectures—the pyramid-shaped BVGN and the isotropic BVGN. The isotropic BVGN was used for interpretability analyses, while the pyramid-shaped BVGN was used to enhance overall performance. Both architectures were initiated with an identical discretization module featuring the same number of nodes. However, as the network deepened, the isotropic BVGN maintained a constant node count, ensuring uniform information flow akin to the transformer model's approach. In contrast, the pyramid-shaped BVGN underwent a process of node consolidation, progressively reducing the number of nodes to form a pyramidal structure, echoing the hierarchical feature extraction of CNNs in computer vision.

BVGN Model Training

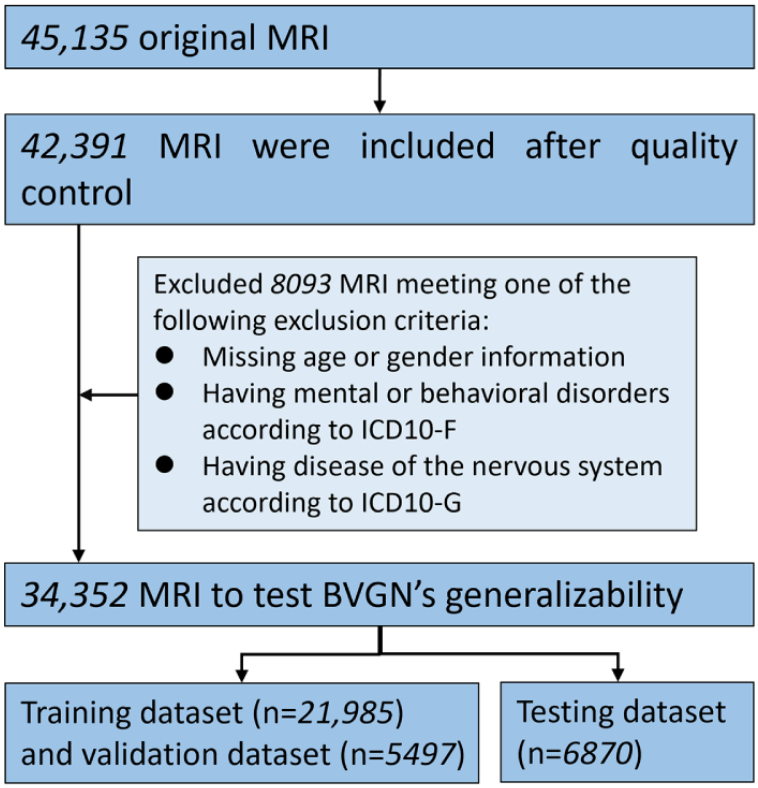
The BVGN model was trained using multiple graphics processing units to ensure efficient computation. For optimization purposes, we implemented the stochastic gradient descent optimizer with an initial learning rate of 0.01, aiming to minimize the Kullback-Leibler divergence between the predicted age distribution and label age distribution. To prevent overfitting, an L2 weight decay coefficient value of 0.001 was applied during training sessions lasting up to 160 epochs. Furthermore, a batch size of 32 was adopted for improved performance, while progressively decreasing the learning rate by multiplying it with a factor of 0.7 after every consecutive span of 30 epochs.

BVGN Model Performance and Generalizability Evaluation

To objectively assess the performance of our framework, we used MAE and Pearson correlation as validation metrics to evaluate the efficacy of BVGN as a prediction model. Furthermore, we simultaneously trained and tested the simple

fully convolutional network (SFCN) and ResNet models on the same dataset, ensuring a fair comparison. To validate the model's generalizability, we retrained it using the UKB dataset, which is large scale and not restricted to disease-specific populations. The detailed inclusion process and demographic characteristics of the included UKB dataset are presented in Figure 3 and Table S1 in Multimedia Appendix 1.

Figure 3. Flow diagram to show inclusion criteria for brain vision graph neural network's (BVGN's) generalizability. ICD: International Classification of Diseases; MRI: magnetic resonance imaging.



BVGN Model Interpretation

In contrast to models exclusively relying on CNNs, BVGN incorporates graph convolutional kernels to modify the relationships among local regions. This enables BVGN to offer interpretability not only through saliency maps but also by considering pertinent properties of the graph structure. More specifically, we first visualized a gradience-based saliency map [45] for participants classified as CN and those with MCI and AD. Subsequently, certified radiologists used the automated anatomical labelling atlas to ascertain the specific brain regions to which the saliency map corresponded. Furthermore, we conceptualized the brain as a graph-based complex system and characterized the interregional cooperative aging patterns by analyzing the connectivity of critical nodes across different hierarchical levels through data-driven visualization techniques.

Cross-Sectional Analysis: Intergroup Comparisons and Logistic Regression

During cross-sectional analysis, each MRI scan was treated as an individual sample, and all statistical analysis was conducted accordingly. Numerical variables in demographics and selection features were presented as mean (SD), while categorical variables were presented as counts (n) and percentages (%). Statistical comparisons between training and testing sets were performed with 2-tailed significance. Specifically, the Student *t* tests and a 1-way ANOVA were used for numerical variables, and chi-square tests were used for categorical variables. A $P<.05$ was considered statistically significant.

For intergroup comparisons, we first characterized the distribution of the brain age gap using the median and IQRs across CN, MCI, and AD classifications. Subsequently, we used the Kruskal-Wallis test followed by 1-tailed Dunn test to evaluate potential significant differences in the brain age gap among these groups, with a significance level set at $P<.05$. Potential confounding effects were addressed through a multiple linear regression model incorporating age and gender as covariates.

We used multiple logistic regression models aiming at binary classification between CN and MCI samples. Our objective was to validate the diagnostic ability of the brain age gap compared to other conventional variables and to assess the potential clinical value of achieving lower MAEs in brain age gap estimation. Inspired by previous studies, we considered the brain age gap along with other relevant cognitive variables, including demographic characteristics, general cognitive assessments, brain volume features, and APOE4 carriage. Given the simplicity and high discriminative and predictive power of FAQ as reported by previous studies, we specifically investigated the impact of combining the brain age gap with FAQ to determine if their integration improves diagnostic classification.

Model performance was evaluated using metrics, including accuracy, precision, recall, F_1-score, the receiver operating characteristic curves, and the area under the receiver operating characteristic curves (AUCs).

We initially established a benchmark model incorporating only demographic variables. Subsequently, brain age gap, various cognitive assessment scores, brain volume characteristics, and APOE4 carriage were added to the benchmark model to construct various logistic regression models. A comprehensive model integrating all the abovementioned variables was then developed, and their coefficients were compared. These analyses aimed to provide a holistic view in comparing the brain age gap with conventionally recognized cognition-related variables. Furthermore, we included both the FAQ and the brain age gap in the benchmark model to explore their combined utility. We also investigated the potential clinical value of lower MAEs by including diverse brain age gaps estimated from models with different precision, both in conjunction with the benchmark and comprehensive models, and compared their performances.

Longitudinal Analysis: Brain Age Gap Trajectories and Risk Quantification

We conducted a longitudinal analysis to investigate changes in brain age gap across follow-up scans of the same participants and assess the risk of brain age gap for cognitive decline. We used BVGN, SFCN, and ResNet models to estimate brain age gap across the total 2392 scans and depicted how the estimated brain age gap changed over follow-up time in different progression groups.

After excluding participants without follow-up records, we performed Cox regression using the baseline brain age gap for both the CN and MCI groups. This allowed us to evaluate the association between brain age gap and cognitive decline, as well as to quantify its hazard ratio (HR). We initially estimated the HR of brain age gap using a univariate Cox proportional hazards model, followed by adjustment for demographic features to control for potential confounding bias. Subsequently, a multivariable Cox proportional hazards model that included all selected variables was used to assess the HR of the brain age gap from a holistic view. In addition, we plotted Kaplan-Meier survival curves stratified by the upper and lower 50% levels of brain age gap in both CN and MCI groups to visualize the performance of the brain age gap in predicting cognitive decline.

Ethical Considerations

Data used in the preparation of this paper were obtained from the ADNI database. Consequently, the investigators within the ADNI contributed to the design and implementation of ADNI and provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found [46] on the internet.

Data collection and sharing for this project were funded by the ADNI (National Institutes of Health grant U01 AG024904) and DOD ADNI (Department of Defense grant W81XWH-12-2-0012).

The data used in the preparation of this paper were obtained from the UKB database. UKB has generic ethical approval from the North West Multi-centre Research Ethics Committee as a Research Tissue Bank (91486), and therefore researchers do not require separate ethical clearance to use the resource.

Results

Demographics and Selection Variable Characteristics

During the BVGN development process, the sample sizes for training, validating, and testing sets were 1228, 308, and 384, respectively. As shown in Table S2 in Multimedia Appendix 1, the mean ages for the training with validating sets and testing sets were 75.65 (SD 6.31) and 75.89 (SD 6.60) years, respectively. The percentages of male individuals in the training and validating set and testing set were 47.1% (724/1536) and 43.5% (167/384), respectively. No significant difference (P=.51) was observed.

Table 1 details the basic information of the dataset for intergroup comparison, which included 4245 MRI scans from 1349 participants. The discrepancy in participant count (1585 in Table 1) reflects instances where participants exhibited multiple cognitive manifestations during the follow-up period, leading to some being counted multiple times across groups. Specifically, the dataset comprised 1500 MRI scans from 499 participants classified as CN, 2014 scans from 709 participants with MCI, and 731 scans from 377 participants experiencing AD. The mean age was 75.39 years for participants classified as CN, 72.96 years for participants with MCI, and 74.76 years for participants with AD. The percentage of male participants was 46.1% (691/1500) for those classified as CN, 55.2% (1112/2014) for those with MCI, and 49.3% (360/731) for those with AD. Significant differences were observed between the CN, MCI, and AD groups for both age and gender characteristics.

Table 1. Demographic characteristics for the dataset of group comparisons.

	CN[a]	MCI[b]	AD[c]	*P* value
MRI[d] scans, n	1500	2014	731	—[e]
Participants[f], n	499	709	377	—
Age (y), mean (SD)	75.39 (6.06)	72.96 (7.45)	74.76 (7.53)	<.001
Gender, n (%)				<.001
Male	691 (46.1)	1112 (55.2)	360 (49.3)	
Female	809 (53.9)	902 (44.8)	371 (50.7)	

[a]CN: cognitively normal.

[b]MCI: mild cognitive impairment.

[c]AD: Alzheimer disease.

[d]MRI: magnetic resonance imaging.

[e]Not applicable.

[f]The total number of participants (n=1585) in Table 1 included instances where some participants had multiple cognitive manifestations during follow-up, leading to duplicates between groups. This total differed from the 1349 unique participants reported in Figure 1. Statistical analysis was performed by ANOVA test for age and chi-square test for gender. Two-tailed significance was used, with $P<.05$ indicating the presence of significance.

For the logistic regression modeling process, 3514 scans were included, with 2635 (74.99%) in the training set and 879 (25.01%) in the testing set. The characteristics of all variables used in the logistic regression are presented in Table 2. No significant differences were observed between training and testing sets for all the selected variables.

Table 2. Characteristics of selected variables in the training and testing datasets for logistic regression.

Variables	Total dataset (n=3514)	Training dataset (n=2635)	Testing dataset (n=879)	*P* value
Education (y), mean (SD)	16.10 (2.81)	16.09 (2.82)	16.12 (2.79)	.77
ADAS13[a] (points), mean (SD)	12.50 (7.07)	12.56 (7.12)	12.32 (6.94)	.37
MMSE[b] (points), mean (SD)	28.22 (1.96)	28.23 (1.97)	28.19 (1.93)	.69
RAVLT_Imm[c] (points), mean (SD)	39.48 (11.99)	39.46 (12.04)	39.54 (11.85)	.86
RAVLT_Learn[d], (points), mean (SD)	4.89 (2.66)	4.87 (2.67)	4.97 (2.63)	.34
RAVLT_Forget[e] (points), mean (SD)	4.17 (2.69)	4.17 (2.67)	4.18 (2.75)	.96
RAVLT_VF%[f] (points), mean (SD)	48.97 (35.42)	49.09 (35.78)	48.60 (34.33)	.72
TRABSCOR[g] (points), mean (SD)	95.92 (56.17)	95.31 (55.33)	97.72 (58.60)	.27
FAQ[h] (points), mean (SD)	1.90 (3.50)	1.87 (3.45)	1.98 (3.65)	.41
Ventricles (cm^3), mean (SD)	36.39 (19.47)	36.48 (19.38)	36.10 (19.75)	.62
Hippocampus (cm^3), mean (SD)	6.99 (1.10)	6.98 (1.10)	7.02 (1.09)	.36
Entorhinal (cm^3), mean (SD)	3.62 (0.73)	3.60 (0.73)	3.65 (0.75)	.11
Fusiform (cm^3), mean (SD)	17.60 (2.59)	17.58 (2.56)	17.69 (2.69)	.28
MidTemp (cm^3), mean (SD)	19.70 (2.79)	19.71 (2.76)	19.68 (2.88)	.77
ICV[i] (cm^3), mean (SD)	1524.19 (158.67)	1523.49 (157.98)	1526.28 (160.80)	.65
BVGN_BAG[j]	3.74 (5.22)	3.79 (5.39)	3.58 (4.68)	.30
SFCN_BAG[k]	3.71 (5.10)	3.73 (5.24)	3.66 (4.67)	.70
ResNer_BAG[l]	4.24 (5.23)	4.30 (5.37)	4.08 (4.76)	.30
Gender, n (%)				.51
Male	1803 (51.3)	1343 (51.0)	460 (52.3)	
Female	1711 (48.7)	1292 (49.0)	419 (47.7)	
Marriage, n (%)				.71
Married	2628 (74.8)	1980 (75.1)	648 (73.7)	
Widowed	400 (11.4)	289 (11.0)	111 (12.6)	
Divorced	338 (9.6)	256 (9.7)	82 (9.3)	
Never married	128 (3.6)	96 (3.6)	32 (3.6)	
Unknown	20 (0.6)	14 (0.5)	6 (0.7)	
APOE4_ε4[m], n (%)				.28
0	2100 (59.8)	1594 (60.5)	506 (57.6)	
1	1181 (33.6)	867 (32.9)	314 (35.7)	
2	233 (6.6)	174 (6.6)	59 (6.7)	
Cognitive states, n (%)				.99
CN[n]	1500 (42.7)	1125 (42.7)	375 (42.7)	
MCI[o]	2014 (57.3)	1510 (57.3)	504 (57.3)	

[a]ADAS13: Alzheimer's Disease Assessment—Cognitive Subscale 13-item version.

[b]MMSE: Mini-Mental State Examination.

[c]RAVLT_Imm: Rey Auditory Verbal Learning Test_immediate.

[d]RAVLT_Learn: Rey Auditory Verbal Learning Test_learning.

[e]RAVLT_Forget: Rey Auditory Verbal Learning Test_forgetting.

[f]RAVLT_VF%: Rey Auditory Verbal Learning Test_percentage_forgetting.

[g]TRABSCOR: trail making test part B, time (to complete in Neuropsychological Battery assessment).

[h]FAQ: Functional Activities Questionnaire.

[i]ICV: intracranial volume.

[j]BVGN_BAG: brain age gap estimated by the brain vision graph neural network model.

[k]SFCN_BAG: brain age gap estimated by the simple fully convolutional network model.

[l]ResNet_BAG: brain age gap estimated by the residual neural network model.

[m]APOE4_ε4: number of apolipoprotein E ε4 alleles.

[n]CN: cognitively normal.

[o]MCI: mild cognitive impairment.

Performance and Generalizability of BVGN

We have simultaneously trained various strong baseline models alongside our proposed model using the same training set and explored their performance on the same test set. The results are shown in Table 3.

Table 3. Performance of benchmark models and the brain vision graph neural network (BVGN) model.

Models	Performance	
	MAE[a] (y)	Pearson correlation coefficient (*r*)
3D ResNet15[b]	3.03	0.76
SFCN[c]	2.41	0.87
Isotropic BVGN	3.08	0.79
Pyramid BVGN	2.39	0.86
Pyramid BVGN (external validation)	2.49	0.92

[a]MAE: mean absolute error.

[b]ResNet15: residual neural network with 15 layers.

[c]SFCN: simple fully convolutional network.

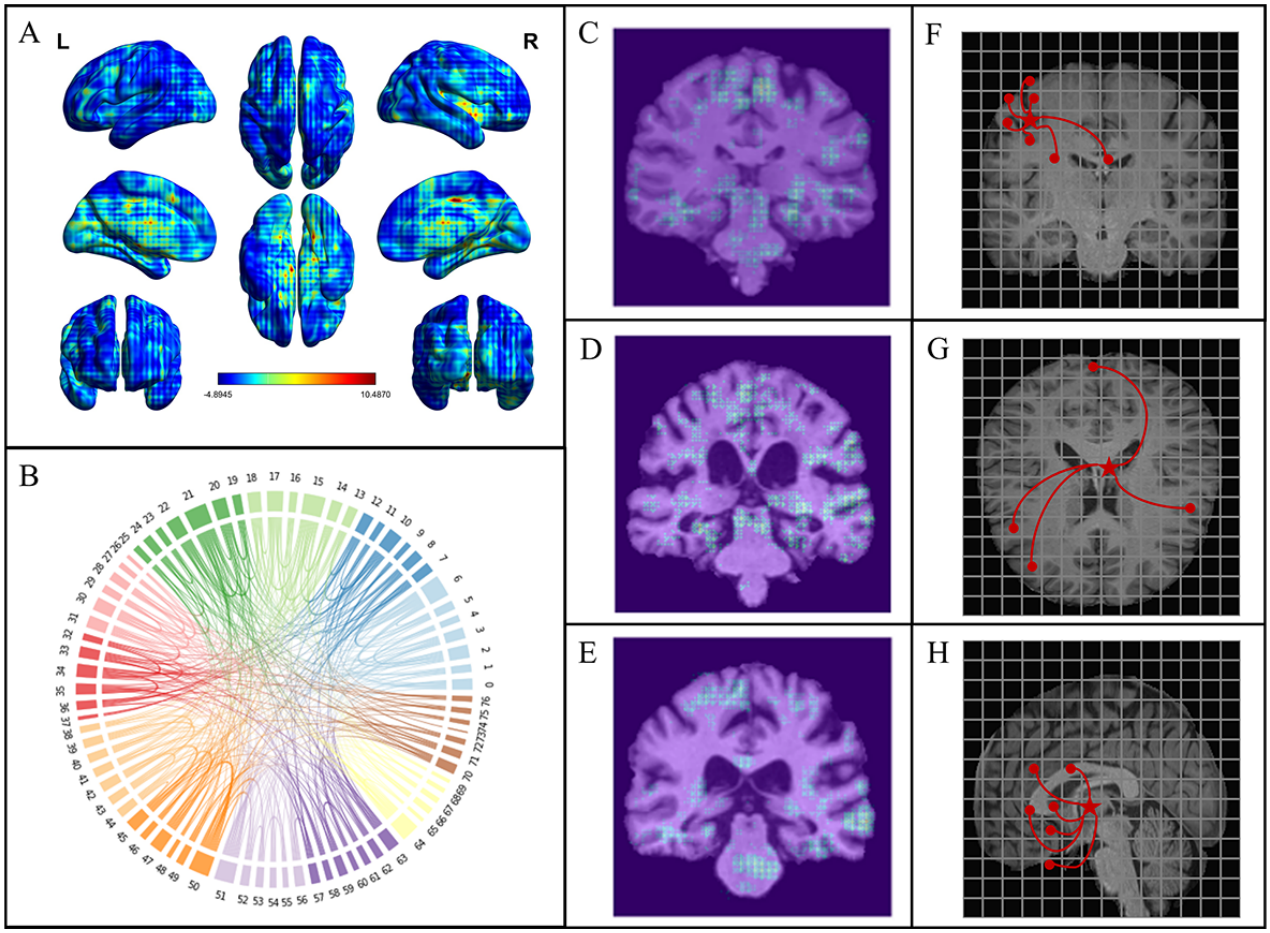
Our isotropic BVGN attained an MAE of 3.08 years, which was slightly worse than the currently optimal model, SFCN, specifically designed for the brain age gap estimation (with an MAE of 2.41 years). After applying deformable convolutions, our pyramid BVGN achieved the best performance among all the tested models with an MAE of 2.39 years and a Pearson correlation coefficient of 0.86 between the predicted and chronological age.

The BVGN model, when retrained on the UKB dataset, exhibited robust performance with an MAE of 2.49 years and a Pearson correlation coefficient of 0.92 (Figure S1 in Multimedia Appendix 1), underscoring its strong generalization capacity.

Interpretability Analysis of BVGN

As shown in Figures 4A and 4B, the saliency map indicated that our model exhibited a moderate focus on various dispersed regions within the cortical layer, while notably emphasizing relatively concentrated parenchymal regions. The regions of concern of the model are different in the population identified as CN (Figure 4C) and those with MCI (Figure 4D) and AD (Figure 4E).

Figure 4. Results for interpretability analysis.

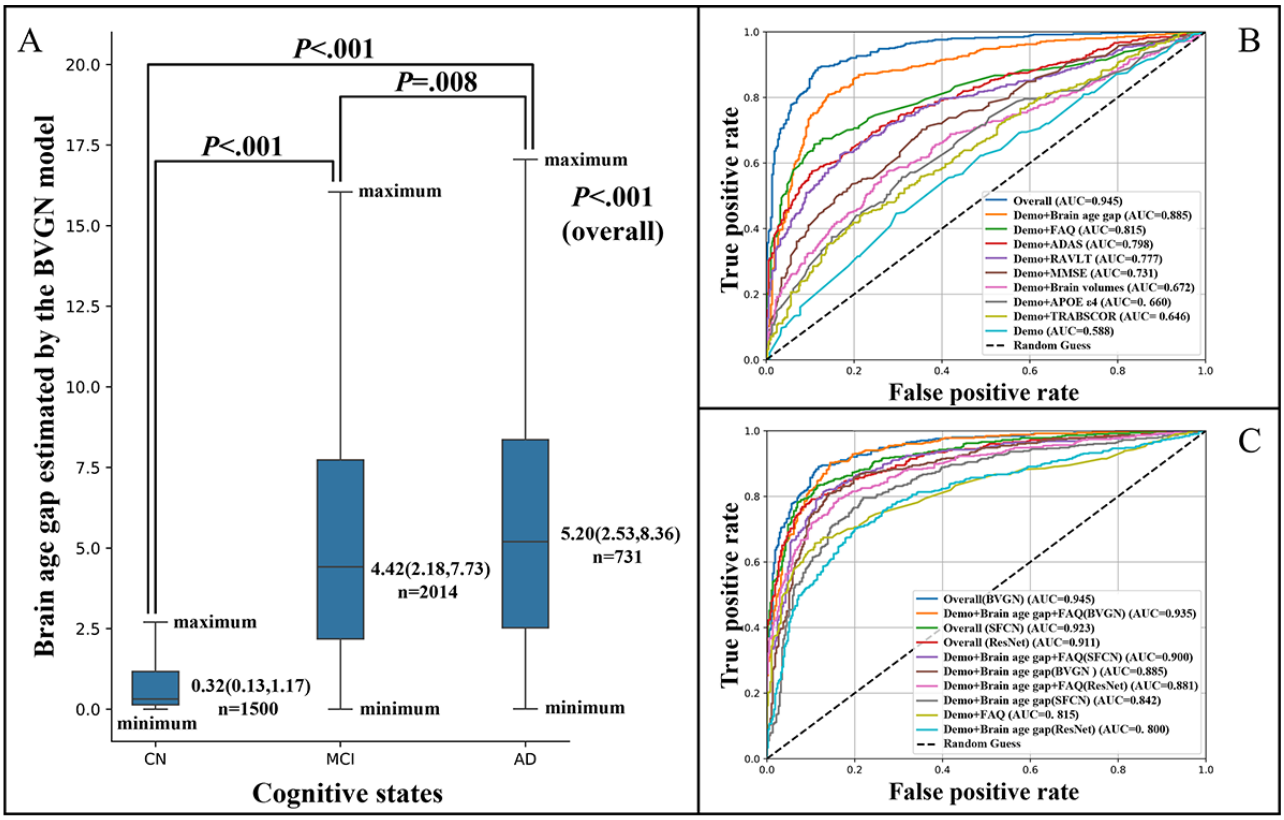


In addition, we described the brain as a complex system by extracting edge and node features from the last layer of an isotropic BVGN backbone. In Figure 4B, the top nodes with the highest degree centrality in the final layer of the isotropic BVGN are depicted, showcasing both these central nodes and their immediate neighboring nodes. This visualization provides insights into the most influential regions within the brain's voxel network as determined by the BVGN's analysis. Specifically, key cortical areas involved the frontal lobe and temporo-parieto-occipital regions, and parenchymal regions included the corpus callosum, cingulate gyrus, and parahippocampal gyrus, covering the striatum, thalamus, and adjacent hippocampus. In Figures 4F-4H, we leveraged a distinctive visualization approach inherent to the vision GNN backbone to illustrate the spatiotemporal dynamics of interregional aging patterns. The red star marks a specific voxel block, and the most relevant regions are delineated through red line connections.

Comparison of Brain Age Gap Among Different Cognitive State Groups

The median of brain age gap was 0.32 (IQR 0.13-1.17) among 1500 CN samples, 4.42 (IQR 2.18-7.73) across 2014 samples with MCI, and 5.20 (IQR 2.53-8.36) in 731 samples with AD (Figure 5A). A statistically significant difference in brain age gap was observed between AD and MCI (P=.008), with MCI demonstrating a significantly higher gap than CN (P<.001), indicating the potential of our model to better distinguish between individuals classified as CN and those with MCI. The results adjusted for age and gender are provided in Table S3 in Multimedia Appendix 1, which were similar to the results mentioned earlier.

Figure 5. Results or group comparisons and receiver operating characteristic curve of multiple logistic regressions. ADAS: Alzheimer's Disease Assessment Scale; APOE4: apolipoprotein E 4 alleles; AUC: area under the receiver operating characteristic curve; BVGN: brain vision graph neural network; CN: cognitively normal; FAQ: Functional Activities Questionnaire; MCI: mild cognitive impairment; MMSE: Mini-Mental State Examination; RAVLT: Rey Auditory Verbal Learning Test; ResNet: residual neural network; SFCN: simple fully convolutional network; TRABSCOR: trail making test part B, time.



Multiple Logistic Regression Models

Brain Age Gap Compared to Other Cognition-Related Variables

Detailed metrics of various logistic regression analyses aiming to compare brain age gap and other cognition-related variables are provided in Table 4, while multiple receiver operating characteristic curves and corresponding AUC scores are illustrated in Figure 5B. Initially, the benchmark model, including only demographic features, yielded an AUC of 0.588. Subsequent incorporation of additional variables, such as FAQ [47], Alzheimer's Disease Assessment Scale-Cognitive Subscale13-item version [48], Rey Auditory Verbal Learning Test [49], Mini-Mental State Examination [50], trail making test part B, time [51], brain volume features, APOE4 carriage, and brain age gap led to improved AUC scores, reaching 0.815, 0.798, 0.777, 0.731, 0.646, 0.672, 0.660, and 0.885, respectively. Notably, the comprehensive model encompassing all variables achieved an AUC of 0.945, with the brain age gap exhibiting the highest coefficient (Figure 6B). These results underscore the robust and superior discriminative ability of the brain age gap in distinguishing between CN and MCI samples.

Table 4. Performance of the logistics regression models in classifying cognitive states: cognitively normal and mild cognitive impairment (brain age gap versus conventional cognitive variables).

Input features	Accuracy	Precision	Recall	F_1-score	AUC[a]
Overall	0.882	0.915	0.875	0.895	0.945
Demographic+BVGN_BAG[b]	0.820	0.890	0.784	0.833	0.885
Demographic+SFCN_BAG[c]	0.780	0.839	0.764	0.800	0.842
Demographic+ResNet_BAG[d]	0.743	0.823	0.702	0.758	0.800
Demographic+FAQ[e]	0.758	0.880	0.669	0.760	0.815
Demographic+ADAS[f]	0.721	0.774	0.726	0.749	0.798
Demographic+RAVLT[g]	0.714	0.759	0.736	0.747	0.777
Demographic+MMSE[h]	0.678	0.723	0.710	0.717	0.731
Demographic+brain volumes	0.617	0.646	0.734	0.687	0.672
Demographic+APOE4[i]	0.631	0.645	0.794	0.712	0.660
Demographic+TRABSCOR[j]	0.611	0.644	0.720	0.680	0.646
Demographic	0.577	0.583	0.919	0.713	0.588

[a]AUC: area under the receiver operating characteristic curve.

[b]BVGN_BAG: brain age gap estimated by the brain vision graph neural network model.

[c]SFCN_BAG: brain age gap estimated by the simple fully convolutional network model.

[d]ResNet_BAG: brain age gap estimated by the residual neural network model.

[e]FAQ: Functional Activities Questionnaire.

[f]ADAS: Alzheimer's Disease Assessment Scale.

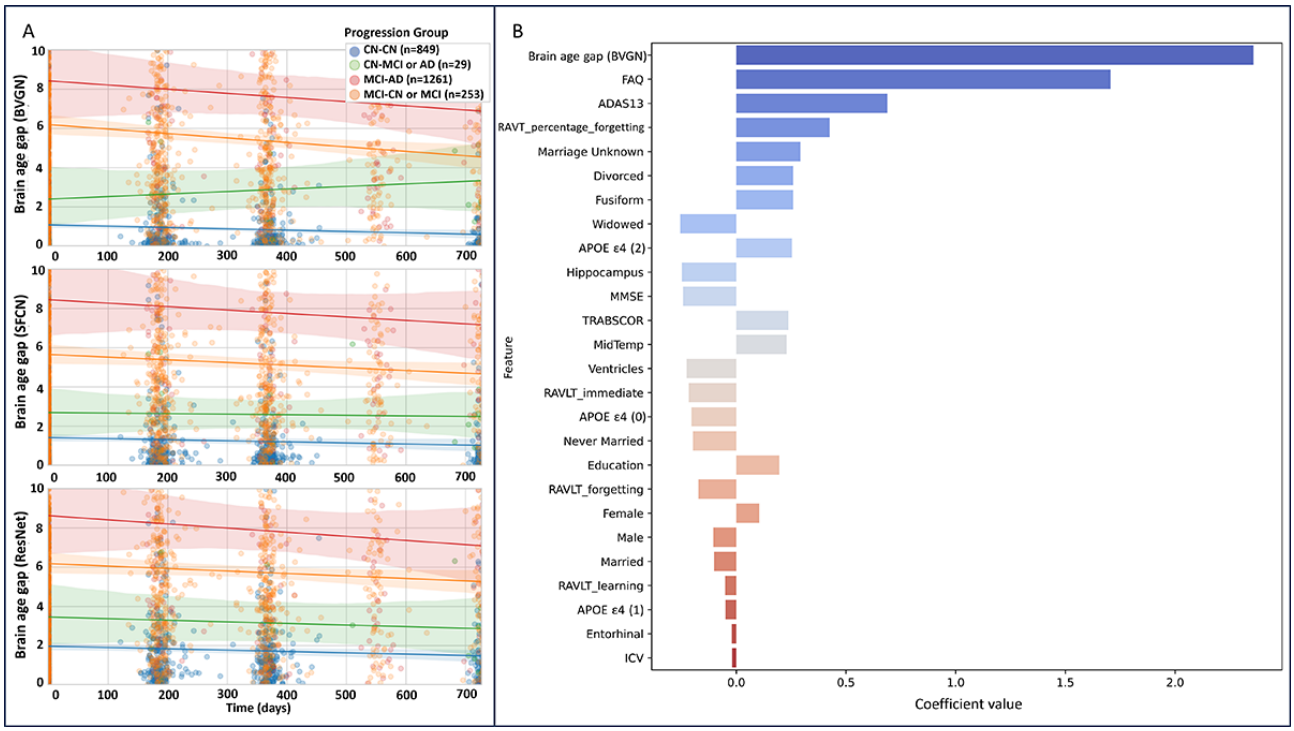
[g]RAVLT: Rey Auditory Verbal Learning Test.

[h]MMSE: Mini-Mental State Examination.

[i]APOE4: apolipoprotein E ε4 alleles.

[j]TRABSCOR: Trail Making Test Part B Time.

Figure 6. (A) Visualization of longitudinal analysis for brain age gap from various models by progression groups over 2 years and (B) coefficients of various variables in the brain vision graph neural network (BVGN) comprehensive logistics regression model discriminating between cognitively normal (CN) and mild cognitive impairment (MCI). ADAS: Alzheimer's Disease Assessment Scale; APOE ε4: apolipoprotein E 4 alleles; AUC: area under the receiver operating characteristic curve; FAQ: Functional Activities Questionnaire; ICV: intracranial volume; MMSE: Mini-Mental State Examination; RAVLT: Rey Auditory Verbal Learning Test; ResNet: residual neural network; SFCN: simple fully convolutional network; TRABSCOR: trail making test part B, time.



Combined Performance of Brain Age Gap With FAQ

In the comprehensive logistics regression model of BVGN, FAQ demonstrated the second-largest coefficient following brain age gap (Figure 6B). When comparing the comprehensive models that included all selective variables with another model incorporating demographic variables, brain age gap and FAQ, minimal declines in AUCs were observed—BVGN decreased by 0.011 (from 0.945 to 0.935), SFCN by 0.024 (from 0.923 to 0.900), and ResNet by 0.031 (from 0.911 to 0.881; Table 5; Figure 5C). Notably, BVGN exhibited the smallest AUC reduction when all variables except FAQ were missing, outperforming other frameworks. In addition, with only demographic variables, brain age gap, and FAQ, BVGN could provide a higher AUC (0.935) compared to other frameworks, even when all selective features were incorporated (SFCN: AUC=0.923; ResNet: AUC=0.911). This underscored the combined efficiency of FAQ and brain age gap, as well as the robustness of BVGN in scenarios with missing variables.

Table 5. Performance of logistics regression models: brain age gaps evaluated by models with different precisions (mean absolute errors) and combined utility of brain age gap with Functional Activity Questionnaire (FAQ).

Input features and model	Accuracy	Precision	Recall	F_1-score	AUC[a]
Overall					
BVGN[b]	0.882	0.915	0.875	0.895	0.945
SFCN[c]	0.851	0.889	0.845	0.867	0.923
ResNet[d]	0.830	0.877	0.819	0.847	0.911
Demographic+brain age gap+FAQ					
BVGN	0.856	0.907	0.833	0.869	0.935
SFCN	0.835	0.895	0.808	0.849	0.900
ResNet	0.807	0.873	0.776	0.821	0.881
Demographic+brain age gap					
BVGN	0.820	0.890	0.784	0.833	0.885
SFCN	0.780	0.839	0.764	0.800	0.842
ResNet	0.743	0.823	0.702	0.758	0.800

[a]AUC: area under the receiver operating characteristic curve.

[b]BVGN: brain vision graph neural network.

[c]SFCN: simple fully convolutional network.

[d]ResNet: residual neural network.

Potential Clinical Value of Lower MAEs in Brain Age Gap Estimation

In assessing the potential clinical utility of models with lower MAEs, our BVGN model consistently outperformed SFCN and ResNet models (Table 5), both in the simplified model comprising only demographic features and brain age gap and in the comprehensive model incorporating all variables. Notably, brain age gap estimated by BVGN demonstrated the highest AUCs (Figure 5C). Specifically, in the simplified model, BVGN achieved AUCs of 0.885, while SFCN and ResNet yielded AUCs of 0.842 and 0.800, respectively. Similarly, in the comprehensive model, BVGN achieved AUCs of 0.945, surpassing SFCN and ResNet with AUCs of 0.923 and 0.911, respectively. These findings were consistent with the ranking of MAEs obtained from these 3 frameworks, indicating that models with lower MAEs might have the potential to enhance the model's discriminatory ability.

Longitudinal Value of BVGN

Brain Age Gap Trajectories in 2 Years

The number of MRI scans for CN-CN, CN-MCI or CN-AD, MCI-CN or MCI-MCI, and MCI-AD were 849, 29, 1261, and 253, respectively.

As illustrated in Figure 6A, we observed consistent ordering of brain age gaps across progression groups in all the frameworks over the 2-year period (MCI-AD>MCI-CN or MCI-MCI>CN-MCI or CN-AD>CN-CN). Moreover, brain age gaps tended to decrease over the 2-year follow-up period in each progression group in both the SFCN and ResNet models, as well as in all groups in the BVGN model except the CN-MCI or CN-AD group. However, in the BVGN model, an interesting increase in the brain age gap was observed in the CN-MCI or CN-AD group.

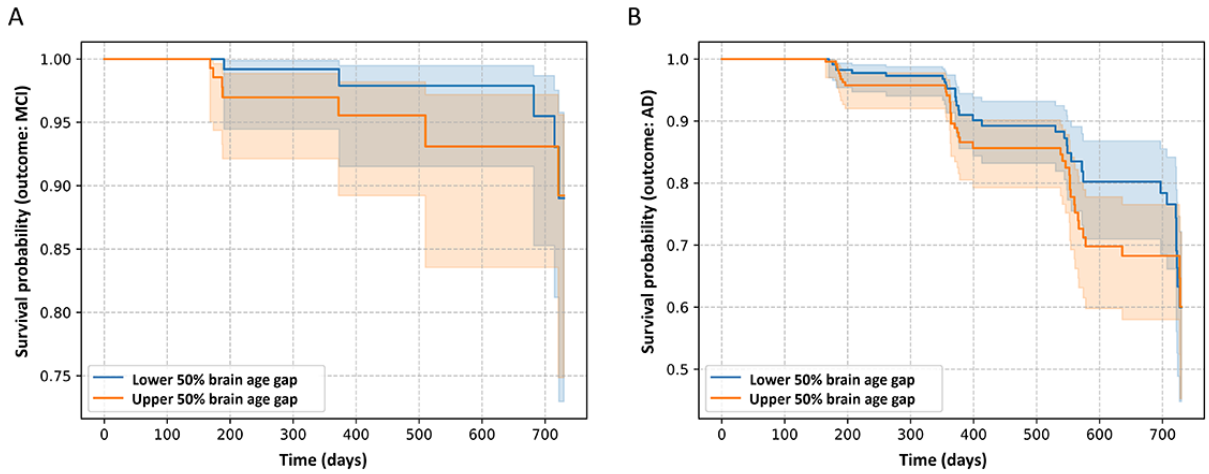
Risk Evaluation by Brain Age Gap

Among the 281 individuals classified as CN at the baseline, 12 (4.3%) declined to MCI, while 269 (95.7%) remained in the CN state. Among the 468 participants with MCI at the baseline, 74 (15.8%) declined to AD while 394 (84.2%) remained in the MCI state or improved to the CN state.

Both univariable Cox regression models and demographic-adjusted as well as fully adjusted multivariable models demonstrated that the brain age gap was a significant risk factor for cognitive decline in both the CN and MCI groups (Table S4 in Multimedia Appendix 1). In the univariable model, the HR for brain age gap was 1.55 (95% CI 1.13-2.13) in the CN group and 1.29 (95% CI 1.09-1.51) in the MCI group. For the CN group, the HR for the brain age gap was 1.47 (95% CI 1.06-2.05) in the demographic-adjusted model and increased to 2.13 (95% CI 1.01-4.47) in the fully adjusted model. In the MCI group, the HR was 1.30 (95% CI 1.10-1.53) for the demographic-adjusted model and 1.24 (95% CI 1.03-1.48) for the fully adjusted model.

The Kaplan-Meier curves indicated that in both CN (Figure 7A) and MCI (Figure 7B) groups, the survival curve for the lower 50% of the brain age gap group was higher than that for the upper 50% group. This finding aligned with the HRs, suggesting that a higher level of brain age gap was associated with an increased risk of cognitive decline.

Figure 7. Visualization of Kaplan-Meier curves stratified by the upper and lower 50% levels of brain age gap in both (A) cognitively normal and (B) mild cognitive impairment (MCI) groups. AD; Alzheimer disease.



Discussion

Principal Findings

In summary, this study proposed a novel framework, BVGN, for brain age estimation. This model not only introduced an innovative modeling approach but also achieved lower MAEs (2.39 years) compared to current state-of-the-art frameworks for brain age gap estimation. It also demonstrated a robust performance on a large-scale external dataset, maintaining strong generalizability (MAE=2.49 years; *r*=0.92; Figure S1 in Multimedia Appendix 1). In addition, BVGN offers an interpretable analysis approach not found in other DL models, which reveals patterns of coaging between brain regions by visualizing the locations of neighboring nodes associated with high-importance nodes. In our study, BVGN provided interpretable attention regions corroborated by medical evidence. Our validation experiments demonstrated its robust performance in MCI diagnosis, its capacity as a risk biomarker for cognitive decline, and its potential to aid in screening and clinical applications.

BVGN integrates deformable convolutional kernels and a dynamic graph for brain age gap estimation. To the best of our knowledge, this work represents the first successful application of vision GNN to T1-weighted MRI data without requiring additional brain segmentation, achieving competitive performance. Unlike 3D CNNs, which process brain images as regular voxel grids and struggle with the inherently irregular morphological characteristics of cerebral regions, often resulting in redundant computations and information loss, BVGN leverages graph-structured representations to preserve complex anatomical features. The integration of deformable convolutional kernels enables the model to focus on specific pathological patterns while suppressing extraneous noise, thereby enhancing local feature modeling [34]. Furthermore, the topological properties of graph structures facilitate multiregional feature aggregation across distant brain regions [38,39], a capability unattainable by grid-based CNNs. To strengthen age prediction accuracy, we transformed age from a discrete numerical variable into a smoothed probability distribution, preserving its continuity and capturing interval similarity [40,41,49,50]. Comparative evaluations demonstrate superior performance of BVGN over existing CNN-based models such as BrainAgeNeXt, achieving an MAE of 2.39 versus 2.78 years, and maintaining robust generalization capabilities (MAE=2.49 years; *r*=0.92) on the UKB dataset after retraining. Structure-functional coupling has garnered increasing attention in neuroscience, with graph-based modeling serving as a critical analytical framework [52]. For instance, structural connectivity derived from diffusion-weighted imaging via white matter fiber tractography and functional connectivity computed from functional MRI are commonly represented as adjacency matrices in graph structures. The framework, being graph-based in design, is inherently more scalable and adaptable to the aforementioned MRI modalities compared to other DL approaches, providing an interface with modal-type embedding for early feature fusion [40]. This capability not only enhances multimodal integration but also helps mitigate the semantic gap across different MRI modalities [53].

We found substantial evidence validating the BVGN's focused brain regions associated with cognitive decline, which provided strong interpretability for our model. The frontal lobes exhibited particularly prominent changes within an annual 0.87% reduction in overall cortical volume [54], while the temporo-parieto-occipital junction was intricately involved in high-level human neurological functions, such as language, memory, calculation, and writing [55]. Shape characteristics in the corpus callosum were demonstrated to have potential for distinguishing cognitive deterioration [56,57], and the cingulate gyrus was implicated in the onset of neurodegenerative and psychiatric disorders [58]. In addition, the parahippocampal gyrus integrates signals from the limbic and neocortex, predicting later memory [59], and the fractional anisotropy alterations in the parahippocampal white matter have been observed across AD stages [60,61].

In recent years, despite the substantial academic focus on the brain age gap, questions have persisted regarding its reliability as a clinical tool compared to cognitive assessments for diagnosing cognitive decline. After multiple experiments that incorporated various kinds of input variables to distinguish CN and MCI, our findings revealed that except for the trail making

test part B, time (AUC=0.646), cognitive assessments exhibited AUCs from 0.731 to 0.815, which surpassed both the AOPE4 carriage (AUC=0.660) and brain volume characteristics (AUC=0.672), and were inferior to brain age gap (AUC=0.885). While cognitive assessments provided affordability, the subjective bias among clinical practitioners might impact diagnosis [62], with limitations in sensitivity observed in certain cases, such as mild levels of impairment [32], subjective deficits, and healthy cognitive aging [63]. According to our findings, our model could offer alternative and more objective evaluations for clinical settings, potentially reducing physicians' workload and enhancing convenience.

The FAQ is a 10-item collateral-report scale with 4-point ordinal responses per item, totaling 0 to 30 points, with higher scores indicating greater functional impairment [64]. This makes it highly accessible in clinical settings and convenient for clinical practitioners. Scientific publications have demonstrated its good internal consistency and high discriminative validity for differentiating between CN and MCI, as well as between MCI and very mild AD [65,66]. Its high predictive validity in detecting individuals at risk of progression from MCI to AD and from CN to MCI has also been reported [47]. In our study, we observed minimal reductions in AUCs when comparing models that incorporated all selective variables and models with only demographic variables, brain age gap, and FAQ. Our selected demographic variables, education years, gender, and marriage status, were deliberately straightforward. This simplicity would enhance the feasibility of our proposed model for clinical applications, especially in cases of uncertain MCI diagnosis. By integrating basic demographics, FAQ responses, and T1-weighted MRI scans, our BVGN model could serve as a reliable assistant tool for clinicians in MCI diagnosis.

In our comprehensive model that integrated brain age gap, cognitive assessments, and other relevant variables, we observed the highest AUC of 0.945 in classification tasks. The integration of variables from diverse perspectives might contribute to improving performance. In addition, we observed brain age gap possessed the highest variable coefficient, with each scalarized unit increase associated with a 10.6 times higher risk of MCI. This further supported the brain age gap's capacity in MCI diagnosis.

Furthermore, our analysis revealed a relationship between model performance, as indicated by lower MAEs, and superior performance in discriminating between CN and MCI. Specifically, BVGN, SFCN, and ResNet achieved MAEs of 2.39, 2.41, and 2.65 years, respectively. Their corresponding AUCs in the simplified and comprehensive logistics models for classifying CN and MCI were 0.885, 0.842, and 0.800 and 0.945, 0.923, and 0.911. This finding underscored the importance of striving for minimal MAE in model development for enhanced predictive accuracy in subsequent cognitive decline tasks.

During longitudinal analysis, we observed an increase in the brain age gap derived from BVGN in CN-MCI or CN-AD progression group within 2 years after baseline. This increase was not observed in either SFCN or ResNet models. This group of participants classified as "CN" might be some individuals at high risk who have already experienced early pathological changes related to dementia yet remain classified as "CN" due to the absence of noticeable cognitive impairments. This finding supported BVGN as a sensitive marker for identifying participants at high risk who, despite being classified as "CN," were likely to progress to MCI in a few years. Existing scientific research suggests that MCI can be reversible to a certain extent, whereas AD, which is associated with significant cognitive degeneration, is challenging to reverse once it has progressed [67-69]. Therefore, early identification of these individuals at high risk for MCI is of great significance for better patient prognosis.

Multiple Cox proportional hazard models and Kaplan-Meier curves revealed a clear association between a higher brain age gap and an increased risk of cognitive decline. Specifically, each 1-unit increase in the brain age gap was associated with a 55% higher risk of cognitive decline in the CN group and a 29% increased risk in the MCI group. This finding underscored the brain age gap derived from BVGN as a sensitive risk biomarker for cognitive decline. Notably, the brain age gap appeared more sensitive in the CN groups. This observation aligned with existing research, which indicated the ability of the brain age gap to detect subtle preclinical or early neurodegenerative changes in populations classified as CN [70,71].

In clinical settings, the BVGN-estimated brain age gap serves as a tool for detecting cognitive decline and assessing the efficacy of innovative therapeutic interventions. For example, in the United States, 2 antiamyloid monoclonal antibodies, lecanemab (brand name Leqembi) and aducanumab (brand name Aduhelm) [72,73], have been approved for the treatment of AD. Despite the potential therapeutic benefits, these interventions are not devoid of inherent risks. Patients require vigilant monitoring for potential side effects, such as amyloid-related imaging abnormalities and infusion reactions, particularly at the onset of treatment. Currently, imaging provides limited insight into a patient's therapeutic response. However, BVGN could theoretically offer a novel approach by capturing changes in brain age pre- and posttreatment with monoclonal antibodies. This could provide an objective measure of a patient's response to the medication, enhancing our ability to assess the efficacy of these treatments.

Limitations and Future Directions

The limitation of BVGN lies in its inability to evaluate the performance and interpretability analysis results on larger datasets. In addition, the use of 3D deformable convolutional kernels in BVGN leads to increased computational costs during the training process when compared to other DL models. Our proposed DL-based brain aging framework is specifically designed for the early screening and prediction of neurodegenerative diseases. Epidemiological and neuroimaging studies indicate that structural alterations linked to neurodegenerative processes begin to manifest or accelerate after the age of 50 years [74]. To better capture brain age deviations closely associated with these pathological conditions, we selected individuals aged >45 years as our study population. This age threshold ensures the inclusion of individuals at the critical transition phase where neurodegeneration-related biomarkers become detectable. However, extending this

framework to the entire life span remains essential to comprehensively understand age-related neurodegenerative processes and identify potential early intervention strategies across all life stages.

Despite its limitations, BVGN remains a valuable tool in the study of brain aging, as it surpasses cognitive assessments, brain volume features, and susceptibility genes in classifying participants as CN and those with MCI, and has been demonstrated to be a risk factor for cognitive decline. This underscores the effectiveness of the brain age gap, estimated through individual MRI using BVGN, as an impactful neuroimaging biomarker. Furthermore, BVGN introduces graph theory analysis as a novel component of interpretability for the first time in brain age gap estimation. In the future, we will leverage the advantages of GNNs in brain analysis, and integrate them with other modalities of medical imaging to further enhance our framework and provide a more robust interpretability analysis based on graph theory. Future work will focus on 3 key directions to enhance the clinical utility and scientific depth of our framework. First, we will explore diverse graph convolutional feature aggregation mechanisms to optimize computational efficiency, enabling broader deployment in real-world clinical environments. Second, we aim to integrate multimodal MRI data by converting different imaging sequences (eg, T1-weighted, T2-weighted, and fluid attenuated inversion recovery) into graph-based representations within the BVGN framework. This multimodal integration will not only enhance the accuracy of brain aging quantification but also enable systematic investigation of cross-modal interactions underlying neurodegenerative processes. Finally, we plan to extend our framework across the entire life span to uncover age-specific aging patterns and their associations with disease progression.

Conclusions

This study proposed a novel DL framework, BVGN, for accurate brain age estimation. Our BVGN model used deformable convolutional kernel to capture brain complex morphology and graph theory to model brain topology, which was applicable to the biological nature of brain MRI. BVGN achieved a superior MAE of 2.39 years in the same testing set compared to other methods, with its attention region aligning with medical evidence of cognitive decline. The robust capacity of BVGN-derived brain age gap was validated through both cross-sectional and longitudinal analysis, which outperformed all conventional tools in aiding the diagnosis of MCI and was sensitive for identifying individuals at high risk for future MCI. The BVGN model can precisely evaluate brain aging and predict cognitive decline, offering substantial potential for improving early identification of neurodegenerative disorders and enhancing clinical applications.

Acknowledgments

This work was supported by the National Natural Science Foundation for Young Scientists of China (grant 82203984), the Healthy Zhejiang One Million People Cohort (grant K-20230085), the Zhejiang Provincial Key Laboratory of Intelligent Preventive Medicine (grant 2020E10004), the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (grant 2019R01007), and the Key Research and Development Program of Zhejiang Province (grant 2020C03002).

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc; Cogstate; Eisai Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health [75]. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The data for external validation was obtained from the UK Biobank database [76].

Authors' Contributions

WL, XW, and ZL contributed to conceptualization, data curation, methodology, funding acquisition, project administration, resources, supervision, Writing, review, and editing. ZL, JL, and JL contributed to data curation, formal analysis, investigation, methodology, validation, visualization, software support, writing the original draft, writing, review, and editing. MW, AX, YH, QY, LZ, and YL contributed to investigation, validation, visualization, writing, review, and editing. MW contributed to data curation and methodology. JB contributed to writing, review, editing, methodology, validation, and resources.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables, figures, and detailed brain vision graph neural network (BVGN) architecture: exhaustive demographics for BVGN development and generalizability validation, Dunn-corrected intergroup comparisons, Cox regression results for brain-age-gap risk, and scatter plot performance across external cohorts.
[DOCX File , 59 KB - aging_v8i1e73004_app1.docx]

References

1. Kanasi E, Ayilavarapu S, Jones J. The aging population: demographics and the biology of aging. Periodontol 2000 2016 Oct;72(1):13-18. [doi: 10.1111/prd.12126] [Medline: 27501488]
2. Kehler DS. Age-related disease burden as a measure of population ageing. Lancet Public Health 2019 Mar;4(3):e123-e124 [FREE Full text] [doi: 10.1016/S2468-2667(19)30026-X] [Medline: 30851865]
3. Ferri CP, Prince M, Brayne C, Brodaty H, Fratiglioni L, Ganguli M, et al. Global prevalence of dementia: a Delphi consensus study. Lancet 2005 Dec 17;366(9503):2112-2117 [FREE Full text] [doi: 10.1016/S0140-6736(05)67889-0] [Medline: 16360788]
4. Grimm A, Eckert A. Brain aging and neurodegeneration: from a mitochondrial point of view. J Neurochem 2017 Nov;143(4):418-431 [FREE Full text] [doi: 10.1111/jnc.14037] [Medline: 28397282]
5. Azam S, Haque ME, Balakrishnan R, Kim IS, Choi DK. The ageing brain: molecular and cellular basis of neurodegeneration. Front Cell Dev Biol 2021;9:683459 [FREE Full text] [doi: 10.3389/fcell.2021.683459] [Medline: 34485280]
6. Ran C, Yang Y, Ye C, Lv H, Ma T. Brain age vector: a measure of brain aging with enhanced neurodegenerative disorder specificity. Hum Brain Mapp 2022 Nov;43(16):5017-5031 [FREE Full text] [doi: 10.1002/hbm.26066] [Medline: 36094058]
7. Muñoz-Ramírez V, Kmetzsch V, Forbes F, Meoni S, Moro E, Dojat M. Subtle anomaly detection: application to brain MRI analysis of de novo Parkinsonian patients. Artif Intell Med 2022 Mar;125:102251 [FREE Full text] [doi: 10.1016/j.artmed.2022.102251] [Medline: 35241258]
8. Lerch JP, van der Kouwe AJ, Raznahan A, Paus T, Johansen-Berg H, Miller KL, et al. Studying neuroanatomy using MRI. Nat Neurosci 2017 Feb 23;20(3):314-326 [FREE Full text] [doi: 10.1038/nn.4501] [Medline: 28230838]
9. Cole JH, Marioni RE, Harris SE, Deary IJ. Brain age and other bodily 'ages': implications for neuropsychiatry. Mol Psychiatry 2019 Feb;24(2):266-281 [FREE Full text] [doi: 10.1038/s41380-018-0098-1] [Medline: 29892055]
10. Cole JH, Franke K. Predicting age using neuroimaging: innovative brain ageing biomarkers. Trends Neurosci 2017 Dec;40(12):681-690 [FREE Full text] [doi: 10.1016/j.tins.2017.10.001] [Medline: 29074032]
11. Rokicki J, Wolfers T, Nordhøy W, Tesli N, Quintana DS, Alnaes D, et al. Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. Hum Brain Mapp 2021 Apr 15;42(6):1714-1726 [FREE Full text] [doi: 10.1002/hbm.25323] [Medline: 33340180]
12. Mishra S, Beheshti I, Khanna P. A review of neuroimaging-driven brain age estimation for identification of brain disorders and health conditions. IEEE Rev Biomed Eng 2023;16:371-385. [doi: 10.1109/RBME.2021.3107372] [Medline: 34428153]
13. Cearns M, Hahn T, Baune BT. Recommendations and future directions for supervised machine learning in psychiatry. Transl Psychiatry 2019 Oct 22;9(1):271 [FREE Full text] [doi: 10.1038/s41398-019-0607-2] [Medline: 31641106]
14. Baecker L, Garcia-Dias R, Vieira S, Scarpazza C, Mechelli A. Machine learning for brain age prediction: introduction to methods and clinical applications. EBioMedicine 2021 Oct;72:103600 [FREE Full text] [doi: 10.1016/j.ebiom.2021.103600] [Medline: 34614461]
15. Lee WH, Antoniades M, Schnack HG, Kahn RS, Frangou S. Brain age prediction in schizophrenia: does the choice of machine learning algorithm matter? Psychiatry Res Neuroimaging 2021 Apr 30;310:111270 [FREE Full text] [doi: 10.1016/j.pscychresns.2021.111270] [Medline: 33714090]
16. Baecker L, Dafflon J, da Costa PF, Garcia-Dias R, Vieira S, Scarpazza C, et al. Brain age prediction: a comparison between machine learning models using region- and voxel-based morphometric data. Hum Brain Mapp 2021 Jun 01;42(8):2332-2346 [FREE Full text] [doi: 10.1002/hbm.25368] [Medline: 33738883]
17. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015 May 28;521(7553):436-444. [doi: 10.1038/nature14539] [Medline: 26017442]
18. Kolbeinsson A, Filippi S, Panagakis Y, Matthews PM, Elliott P, Dehghan A, et al. Accelerated MRI-predicted brain ageing and its associations with cardiometabolic and brain disorders. Sci Rep 2020 Nov 17;10(1):19940 [FREE Full text] [doi: 10.1038/s41598-020-76518-z] [Medline: 33203906]
19. Wood DA, Kafiabadi S, Busaidi AA, Guilhem E, Montvila A, Lynch J, et al. Accurate brain-age models for routine clinical MRI examinations. Neuroimage 2022 Apr 01;249:118871 [FREE Full text] [doi: 10.1016/j.neuroimage.2022.118871] [Medline: 34995797]
20. He S, Feng Y, Grant PE, Ou Y. Deep relation learning for regression and its application to brain age estimation. IEEE Trans Med Imaging 2022 Sep;41(9):2304-2317 [FREE Full text] [doi: 10.1109/TMI.2022.3161739] [Medline: 35320092]
21. F A, La Rosa RF, Sather A, Rechtman E, Jalees M, Nabeel I, et al. MRI signature of brain age underlying post- traumatic stress disorder in World Trade Center responders. medRxiv. Preprint posted online on October 19, 2024 [FREE Full text] [doi: 10.1101/2024.10.18.24315761]

22. Zhang Z, Cui P, Zhu W. Deep learning on graphs: a survey. IEEE Trans Knowl Data Eng 2022 Jan 1;34(1):249-270. [doi: 10.1109/TKDE.2020.2981333]
23. Peng H, Xie P, Liu L, Kuang X, Wang Y, Qu L, et al. Morphological diversity of single neurons in molecularly defined cell types. Nature 2021 Oct 06;598(7879):174-181 [FREE Full text] [doi: 10.1038/s41586-021-03941-1] [Medline: 34616072]
24. Tadić B, Andjelković M, Melnik R. Functional geometry of human connectomes. Sci Rep 2019 Aug 19;9(1):12060 [FREE Full text] [doi: 10.1038/s41598-019-48568-5] [Medline: 31427676]
25. Despotović I, Goossens B, Philips W. MRI segmentation of the human brain: challenges, methods, and applications. Comput Math Methods Med 2015;2015:450341 [FREE Full text] [doi: 10.1155/2015/450341] [Medline: 25945121]
26. Han K, Wang Y, Guo J, Tang Y, Wu E. Vision GNN: an image is worth graph of nodes. In: Proceedings of the 36th International Conference on Neural Information Processing Systems. 2022 Presented at: NIPS'22; November 28-December 9, 2022; New Orleans, LA. [doi: 10.48550/arXiv.2206.00272]
27. Weiner MW, Veitch DP. Introduction to special issue: overview of Alzheimer's disease neuroimaging initiative. Alzheimers Dement 2015 Jul;11(7):730-733 [FREE Full text] [doi: 10.1016/j.jalz.2015.05.007] [Medline: 26194308]
28. Toga AW, Crawford KL. The Alzheimer's disease neuroimaging initiative informatics core: a decade in review. Alzheimers Dement 2015 Jul;11(7):832-839 [FREE Full text] [doi: 10.1016/j.jalz.2015.04.004] [Medline: 26194316]
29. Weiner MW, Kanoria S, Miller MJ, Aisen PS, Beckett LA, Conti C, et al. Overview of Alzheimer's disease neuroimaging initiative and future clinical trials. Alzheimers Dement 2025 Jan;21(1):e14321 [FREE Full text] [doi: 10.1002/alz.14321] [Medline: 39711072]
30. Esteban J, Starr A, Willetts R, Hannah P, Bryanston-Cross P. A Review of data fusion models and architectures: towards engineering guidelines. Neural Comput Appl 2005 Jun 21;14(4):273-281. [doi: 10.1007/s00521-004-0463-7]
31. Jack CRJ, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J Magn Reson Imaging 2008 Apr;27(4):685-691 [FREE Full text] [doi: 10.1002/jmri.21049] [Medline: 18302232]
32. Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA. A review of screening tests for cognitive impairment. J Neurol Neurosurg Psychiatry 2007 Aug;78(8):790-799 [FREE Full text] [doi: 10.1136/jnnp.2006.095414] [Medline: 17178826]
33. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 2016 Nov;19(11):1523-1536 [FREE Full text] [doi: 10.1038/nn.4393] [Medline: 27643430]
34. Mansour L S, Di Biase MA, Smith RE, Zalesky A, Seguin C. Connectomes for 40,000 UK Biobank participants: a multi-modal, multi-scale brain network resource. Neuroimage 2023 Dec 01;283:120407 [FREE Full text] [doi: 10.1016/j.neuroimage.2023.120407] [Medline: 37839728]
35. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL. Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 2011 Jan 01;54(1):313-327 [FREE Full text] [doi: 10.1016/j.neuroimage.2010.07.033] [Medline: 20656036]
36. Yaniv Z, Lowekamp BC, Johnson HJ, Beare R. SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. J Digit Imaging 2018 Jun 27;31(3):290-303 [FREE Full text] [doi: 10.1007/s10278-017-0037-8] [Medline: 29181613]
37. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. Neuroimage 2012 Aug 15;62(2):782-790. [doi: 10.1016/j.neuroimage.2011.09.015] [Medline: 21979382]
38. Dai J, Qi H, Xiong Y, Li Y, Zhang G, Hu H. Deformable convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision. 2017 Presented at: ICCV 2017; October 22-29, 2017; Venice, Italy. [doi: 10.1109/iccv.2017.89]
39. Zhu X, Hu H, Lin S, Dai J. Deformable ConvNets V2: more deformable, better results. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019 Presented at: CVPR 2019; June 15-20, 2019; Long Beach, CA. [doi: 10.1109/cvpr.2019.00953]
40. Kim W, Son B, Kim I. ViLT: vision-and-language transformer without convolution or region supervision. In: Proceedings of the 38th International Conference on Machine Learning. 2021 Presented at: ICML 2021; July 18-21, 2021; Online. [doi: 10.48550/arXiv.2102.03334]
41. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16x16 words: transformers for image recognition at scale. In: Proceedings of the Ninth International Conference on Learning Representations. 2021 Presented at: ICLR 2021; May 3-7, 2021; Online. [doi: 10.48550/arXiv.2010.11929]
42. Li G, Müller M, Thabet A, Ghanem B. DeepGCNs: can GCNs go as deep as CNNs? In: Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019 Presented at: ICCV 2019; October 27-November 02, 2019; Seoul, Korea (South). [doi: 10.1109/iccv.2019.00936]
43. Huo Z, Yang X, Xing C, Zhou Y, Hou P, Lv J. Deep age distribution learning for apparent age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2016 Presented at: CVPRW 2016; June 26-July 01, 2016; Las Vegas, NV. [doi: 10.1109/cvprw.2016.95]

44. Gao BB, Zhou HY, Wu J, Geng X. Age estimation using expectation of label distribution learning. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. 2018 Presented at: IJCAI 2018; July 13-19, 2018; Stockholm, Sweden. [doi: 10.24963/ijcai.2018/99]
45. Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps. In: Proceedings of the International Conference on Learning Representations. 2014 Presented at: ICLR 2014; April 14-16, 2014; Banff, AB. [doi: 10.48550/arXiv.1312.6034]
46. Acknowledgement list for ADNI Publications. Alzheimer's Disease Neuroimaging Initiative. URL: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf [accessed 2025-07-18]
47. Marshall GA, Zoller AS, Lorius N, Amariglio RE, Locascio JJ, Johnson KA, et al. Functional activities questionnaire items that best discriminate and predict progression from clinically normal to mild cognitive impairment. Curr Alzheimer Res 2015;12(5):493-502 [FREE Full text] [doi: 10.2174/1567205012051505261150003] [Medline: 26017560]
48. Skinner J, Carvalho JO, Potter GG, Thames A, Zelinski E, Crane PK, et al. The Alzheimer's Disease Assessment Scale-Cognitive-Plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI. Brain Imaging Behav 2012 Dec 22;6(4):489-501 [FREE Full text] [doi: 10.1007/s11682-012-9166-3] [Medline: 22614326]
49. Schoenberg MR, Dawson KA, Duff K, Patton D, Scott JG, Adams RL. Test performance and classification statistics for the Rey Auditory Verbal Learning Test in selected clinical samples. Arch Clin Neuropsychol 2006 Oct;21(7):693-703. [doi: 10.1016/j.acn.2006.06.010] [Medline: 16987634]
50. Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. Nutrients 2021 Nov 08;13(11):3974 [FREE Full text] [doi: 10.3390/nu13113974] [Medline: 34836228]
51. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Ski 1958 Dec;8(7):271-276. [doi: 10.2466/PMS.8.7.271-276]
52. Fotiadis P, Parkes L, Davis KA, Satterthwaite TD, Shinohara RT, Bassett DS. Structure-function coupling in macroscale human brain networks. Nat Rev Neurosci 2024 Oct;25(10):688-704. [doi: 10.1038/s41583-024-00846-6] [Medline: 39103609]
53. Bao H, Wang W, Dong L, Liu Q, Mohammed OK, Aggarwal K, et al. VLMO: unified vision-language pre-training with mixture-of-modality-experts. In: Proceedings of the 36th Conference on Neural Information Processing Systems. 2022 Presented at: NeurIPS 2022; November 28-December 9, 2022; New Orleans, LA. [doi: 10.48550/arXiv.2111.02358]
54. Cox SR, Harris MA, Ritchie SJ, Buchanan CR, Valdés Hernández MC, Corley J, et al. Three major dimensions of human brain cortical ageing in relation to cognitive decline across the eighth decade of life. Mol Psychiatry 2021 Jun 04;26(6):2651-2662 [FREE Full text] [doi: 10.1038/s41380-020-00975-1] [Medline: 33398085]
55. De Benedictis A, Duffau H, Paradiso B, Grandi E, Balbi S, Granieri E, et al. Anatomo-functional study of the temporo-parieto-occipital region: dissection, tractographic and brain mapping evidence from a neurosurgical perspective. J Anat 2014 Aug;225(2):132-151 [FREE Full text] [doi: 10.1111/joa.12204] [Medline: 24975421]
56. Mangalore S, Mukku SS, Vankayalapati S, Sivakumar PT, Varghese M. Shape profile of corpus callosum as a signature to phenotype different dementia. J Neurosci Rural Pract 2021 Jan;12(1):185-192 [FREE Full text] [doi: 10.1055/s-0040-1716805] [Medline: 33531781]
57. Kamal S, Park I, Kim YJ, Kim YJ, Lee U. Alteration of the corpus callosum in patients with Alzheimer's disease: deep learning-based assessment. PLoS One 2021;16(12):e0259051 [FREE Full text] [doi: 10.1371/journal.pone.0259051] [Medline: 34941878]
58. Foster BL, Koslov SR, Aponik-Gremillion L, Monko ME, Hayden BY, Heilbronner SR. A tripartite view of the posterior cingulate cortex. Nat Rev Neurosci 2023 Mar 01;24(3):173-189 [FREE Full text] [doi: 10.1038/s41583-022-00661-x] [Medline: 36456807]
59. Husain M. The parahippocampal region: organization and role in cognitive function. J Neurol Neurosurg Psychiatry 2003;74:829 [FREE Full text] [doi: 10.1136/jnnp.74.6.829-a]
60. Dalboni da Rocha JL, Bramati I, Coutinho G, Tovar Moll F, Sitaram R. Fractional anisotropy changes in parahippocampal cingulum due to Alzheimer's disease. Sci Rep 2020 Feb 14;10(1):2660 [FREE Full text] [doi: 10.1038/s41598-020-59327-2] [Medline: 32060334]
61. Salat DH, Tuch DS, van der Kouwe AJ, Greve DN, Pappu V, Lee SY, et al. White matter pathology isolates the hippocampal formation in Alzheimer's disease. Neurobiol Aging 2010 Feb;31(2):244-256 [FREE Full text] [doi: 10.1016/j.neurobiolaging.2008.03.013] [Medline: 18455835]
62. McClintock AH, Fainstad T, Jauregui J, Yarris LM. Countering bias in assessment. J Grad Med Educ 2021 Oct;13(5):725-726 [FREE Full text] [doi: 10.4300/JGME-D-21-00722.1] [Medline: 34721804]
63. Hess C, Levy B, Hashmi AZ, Hogan J, Greenspan S, Elber A, et al. Subjective versus objective assessment of cognitive functioning in primary care. J Am Board Fam Med 2020;33(3):417-425 [FREE Full text] [doi: 10.3122/jabfm.2020.03.190265] [Medline: 32430373]
64. Pfeffer RI, Kurosaki TT, Harrah CHJ, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]

65. González DA, Gonzales MM, Resch ZJ, Sullivan AC, Soble JR. Comprehensive evaluation of the functional activities questionnaire (FAQ) and its reliability and validity. Assessment 2022 Jun;29(4):748-763 [FREE Full text] [doi: 10.1177/1073191121991215] [Medline: 33543638]
66. Teng E, Becker BW, Woo E, Knopman DS, Cummings JL, Lu PH. Utility of the functional activities questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer disease. Alzheimer Dis Assoc Disord 2010;24(4):348-353 [FREE Full text] [doi: 10.1097/WAD.0b013e3181e2fc84] [Medline: 20592580]
67. Ahmadzadeh M, Christie GJ, Cosco TD, Arab A, Mansouri M, Wagner KR, et al. Neuroimaging and machine learning for studying the pathways from mild cognitive impairment to Alzheimer's disease: a systematic review. BMC Neurol 2023 Aug 22;23(1):309 [FREE Full text] [doi: 10.1186/s12883-023-03323-2] [Medline: 37608251]
68. Öksüz N, Ghouri R, Taşdelen B, Uludüz D, Özge A. Mild cognitive impairment progression and Alzheimer's disease risk: a comprehensive analysis of 3553 cases over 203 months. J Clin Med 2024 Jan 17;13(2):518 [FREE Full text] [doi: 10.3390/jcm13020518] [Medline: 38256652]
69. Long JM, Holtzman DM. Alzheimer disease: an update on pathobiology and treatment strategies. Cell 2019 Oct 03;179(2):312-339 [FREE Full text] [doi: 10.1016/j.cell.2019.09.001] [Medline: 31564456]
70. Cole JH, Ritchie SJ, Bastin ME, Valdés Hernández MC, Muñoz Maniega S, Royle N, et al. Brain age predicts mortality. Mol Psychiatry 2018 May 25;23(5):1385-1392 [FREE Full text] [doi: 10.1038/mp.2017.62] [Medline: 28439103]
71. Franke K, Gaser C. Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained? Front Neurol 2019;10:789 [FREE Full text] [doi: 10.3389/fneur.2019.00789] [Medline: 31474922]
72. Cummings J. Anti-amyloid monoclonal antibodies are transformative treatments that redefine Alzheimer's disease therapeutics. Drugs 2023 May 15;83(7):569-576 [FREE Full text] [doi: 10.1007/s40265-023-01858-9] [Medline: 37060386]
73. Agarwal A, Gupta V, Brahmbhatt P, Desai A, Vibhute P, Joseph-Mathurin N, et al. Amyloid-related imaging abnormalities in Alzheimer disease treated with anti-amyloid-β therapy. Radiographics 2023 Sep;43(9):e230009 [FREE Full text] [doi: 10.1148/rg.230009] [Medline: 37651273]
74. Yang Z, Wen J, Erus G, Govindarajan ST, Melhem R, Mamourian E, et al. Brain aging patterns in a large and diverse cohort of 49,482 individuals. Nat Med 2024 Oct;30(10):3015-3026 [FREE Full text] [doi: 10.1038/s41591-024-03144-x] [Medline: 39147830]
75. Raising private funds to expand NIH impact. The Foundation for the National Institutes of Health. URL: https://fnih.org/ [accessed 2025-05-29]
76. Health research data for the world. UK Biobank. URL: https://www.ukbiobank.ac.uk/ [accessed 2025-07-18]

Abbreviations

AD: Alzheimer disease
ADNI: Alzheimer’s Disease Neuroimaging Initiative
AUC: area under the receiver operating characteristic curve
BVGN: brain vision graph neural network
CN: cognitively normal
CNN: convolutional neural network
DL: deep learning
FAQ: Functional Activities Questionnaire
GNN: graph neural network
HR: hazard ratio
MAE: mean absolute error
MCI: mild cognitive impairment
MRI: magnetic resonance imaging
ResNet: residual neural network
SFCN: simple fully convolutional network
UKB: UK Biobank

Edited by N Schütz; submitted 23.02.25; peer-reviewed by C Yao, A Invernizzi; comments to author 15.05.25; revised version received 10.06.25; accepted 27.06.25; published 01.08.25.

Please cite as:
Li Z, Li J, Li J, Wang M, Xu A, Huang Y, Yu Q, Zhang L, Li Y, Li Z, Wu X, Bu J, Li W
Development and Validation of a Brain Aging Biomarker in Middle-Aged and Older Adults: Deep Learning Approach
JMIR Aging 2025;8:e73004
URL: https://aging.jmir.org/2025/1/e73004
doi:10.2196/73004
PMID:

Original Paper

Online Communities as a Support System for Alzheimer Disease and Dementia Care: Large-Scale Exploratory Study

Sidharth Kaliappan[1], MS; Chunyu Liu[2], MS; Yoshee Jain[3]; Ravi Karkar[1], PhD; Koustuv Saha[3], PhD

[1]Manning College of Information & Computer Sciences, University of Massachusetts Amherst, Amherst, MA, United States

[2]School of Information Sciences, University of Illinois Urbana-Champaign, Champaign, United States

[3]Siebel School of Computing and Data Science, The Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, IL, United States

Corresponding Author:
Koustuv Saha, PhD
Siebel School of Computing and Data Science
The Grainger College of Engineering
University of Illinois Urbana-Champaign
201 N. Goodwin Ave
Siebel Center for Computer Science, 4212
Urbana, IL, 61801
United States
Phone: 1 2172443824
Email: ksaha2@illinois.edu

Abstract

Background: Alzheimer disease (AD) is the leading type of dementia, demanding comprehensive understanding and intervention strategies. In the United States, where over 6 million people are impacted, the prevalence of AD and related dementias (AD/ADRD) presents a growing public health challenge. However, individuals living with AD/ADRD and their caregivers frequently express feelings of marginalization, describing interactions characterized by perceptions of patient infantilization and a lack of respect.

Objective: This study aimed to address 2 key research questions (RQs). For RQ1, we investigated the needs and concerns expressed by participants in online social communities focused on AD/ADRD, specifically on 2 platforms–Reddit's r/Alzheimers and ALZConnected. For RQ2, we examined the prevalence and distribution of social support corresponding to these needs and concerns, and the association between these needs and received support.

Methods: We collected 13,429 posts and comments from the r/Alzheimers subreddit spanning July 2014 to November 2023, and 90,113 posts and comments from ALZConnected between December 2020 (the community's earliest post) and November 2023. We conducted topic modeling using latent Dirichlet allocation (LDA), followed by labeling to identify the major topical themes of discussions. We used transfer learning classifiers to identify the occurrences of emotional support (ES) and informational support (IS) in the comments (or responses) in the discussions. We built regression models to examine how various topical themes are associated with the kinds of support received.

Results: Our analysis revealed a diverse range of topics reflecting community members' varying needs and concerns of individuals affected by AD/ADRD. These themes encapsulate the primary discussions within the online communities: memory care, nursing and caregiving, gratitude and acknowledgment, and legal and financial considerations. Our findings indicated a higher prevalence of IS compared to ES. Regression models revealed that ES primarily occurs in posts relating to nursing and caring, and IS primarily occurs in posts concerning medical conditions and diagnosis, legal and financial, and caregiving at home.

Conclusions: This study reveals that online communities dedicated to AD/ADRD support engage in discussions on a wide range of topics, such as memory care, nursing, caregiving, and legal and financial challenges. The findings shed light on the key pain points and concerns faced by individuals managing AD/ADRD in their households, revealing how they leverage online platforms for guidance and support. These insights underscore the need for targeted institutional and social interventions to address the specific needs of AD/ADRD patients, caregivers, and other family members.

(JMIR Aging 2025;8:e68890) doi:10.2196/68890

KEYWORDS

social media; natural language; Alzheimer disease; social support; online communities; machine learning

Introduction

Alzheimer disease and related dementias (AD/ADRD) consist of neurodegenerative conditions characterized by cognitive impairment, memory loss, and executive function decline [1]. Over 6 million individuals in the United States grapple with the challenges imposed by AD/ADRD [2], signaling a substantial and escalating public health concern, and underscoring the urgent need for comprehensive understanding and intervention. However, despite the existence of available resources, the long-term services and support for people with AD/ADRD exhibit inconsistencies within the existing health care system [3,4]. People with AD/ADRD and their caregivers often perceive interactions with others as reflective of patient infantilization and a lack of sensitivity and respect [5]. They frequently encounter added challenges, encompassing mental health issues, social exclusion, and instances of discrimination [6,7]. In fact, if caregivers' concerns are overlooked, it may accelerate cognitive decline in individuals with AD/ADRD [8].

As technology gets more integrated into our lives and society, previous research has also noted the use of technology by AD/ADRD individuals and their caregivers [9]. Research highlights the significance of acknowledging and addressing the technology requirements and acceptance within AD/ADRD care [10]. However, more emphasis is placed on incorporating technology into AD/ADRD care, with the objective of improving the user experience to assist both patients and caregivers in managing daily life and monitoring their health status on a regular basis. A significant concern among AD/ADRD caregivers, however, remains the limited availability of adequate support systems at both institutional and social levels [11,12]. With the ubiquity of the internet, more individuals are turning to online spaces to seek information, discuss challenges, and share personal experiences. In particular, online communities are spaces that enhance a sense of belonging by connecting individuals with similar experiences. These online communities enable individuals to candidly self-disclose their concerns and seek support from others [13-20]. Previous work also noted how these online platforms could play a role in shaping the experiences of those living with progressive AD/ADRD [21-23]. These online support communities are very much in line with the social support behavioral theory (SSBC) that identifies the interactions intended to provide social support toward helping individuals cope with stress, enhance well-being, and foster positive relationships [24]. Nonetheless, social expectations and requirements in these online communities are largely unknown. Such knowledge would not only enhance our understanding of how the needs and concerns of AD/ADRD individuals and caregivers evolve over time but also help inform the design of targeted online interventions to address these needs.

This study aims to explore the needs and concerns of members within online communities on AD/ADRD, as well as the social support they actively provide and seek. To avoid any biases embedded within specific communities, we examined 2 distinct online communities—r/Alzheimers on Reddit and AlzConnected—dedicated discussion forums for AD/ADRD. Our study is guided by the following research questions (RQs): (1) RQ1: What are the needs and concerns people discuss in online AD/ADRD communities? (2) RQ2: What is the prevalence and distribution of social support across various needs and concerns, and how are different types of needs associated with the types of support received?

Methods

Data

We sourced our data from 2 distinct online AD/ADRD communities—r/Alzheimers on Reddit and AlzConnected, which we describe below:

Reddit

Reddit is a widely used semianonymous online platform that houses a network of over 52 million active users and 100 thousand active online communities. These communities are called "subreddits,"—each offering demographic, topical, or interest-specific discussion boards. The semianonymity of Reddit is known to enable candid self-disclosure and seek social support on stigmatized and sensitive topics [13,25,26], including discussions related to AD/ADRD [22,23,27].

We used Reddit's search feature to look up subreddits related to AD/ADRD, and 2 of the largest subreddits returned were r/AlzheimersGroup (over 117 thousand members) and r/Alzheimers (over 13 thousand members). We then conducted sanity checks to understand the content of these subreddits. These investigations revealed that r/AlzheimersGroup was a community centered on humor and sarcasm, where people predominantly pretended to have forgetfulness and shared memes and cartoon references to Garfield [28]. r/AlzheimersGroup also self-describes that "Serious discussion belongs in r/Alzheimers," Therefore, we omitted r/AlzheimersGroup from our ensuing analyses. The r/Alzheimers subreddit consisted of discussions where people shared troubles and provided social support on AD/ADRD. We collected all the posts and comments data from this subreddit—1000 posts and 12,429 comments between July 2014 and November 2023.

AlzConnected

AlzConnected is another freely accessible website (alzconnected.org) consisting of discussion boards dedicated to AD/ADRD patients and their caregivers. Although a more recent online community, it has already grown a significant user base and is supported by the Alzheimer Association (alz.org). Here, community members express AD/ADRD-related concerns and provide and seek support from one another. This community has multiple thematic discussion boards—(1) living with AD/ADRD, further broken down into "I am living with Alzheimer or other dementia and I am living with younger-onset Alzheimer," and (2) Supporting someone living with AD/ADRD, further broken down into, "I am a caregiver (general topics), caring for a spouse or partner, caring for a parent, caring long distance, and supporting those who have lost someone."

We adopted web scraping approaches to collect the entire data from ALZConnected amounting to 10,328 posts and 79,785 comments from the community between December 2020 (earliest post in the community) and November 2023.

Table 1 summarizes our dataset from the 2 online communities.

Table 1. Summary of dataset from online communities catering to discussions on AD/ADRD[a].

Community	Number of posts	Number of comments
r/Alzheimers	1000	12,429
ALZConnected	10,328	79,785

[a]AD/ADRD: Alzheimer disease and related dementias.

Analyzing the Needs and Concerns

Toward answering RQ1, we conducted unsupervised topic modeling followed by human annotation to obtain the needs and concerns expressed by members of AD/ADRD online communities.

Topic Modeling and Thematic Analysis

We started by examining the content of posts and comments in r/Alzheimers and AlzConnected to gain insights into the prevalent needs and concerns discussed within AD/ADRD online communities.

Building a Topic Model

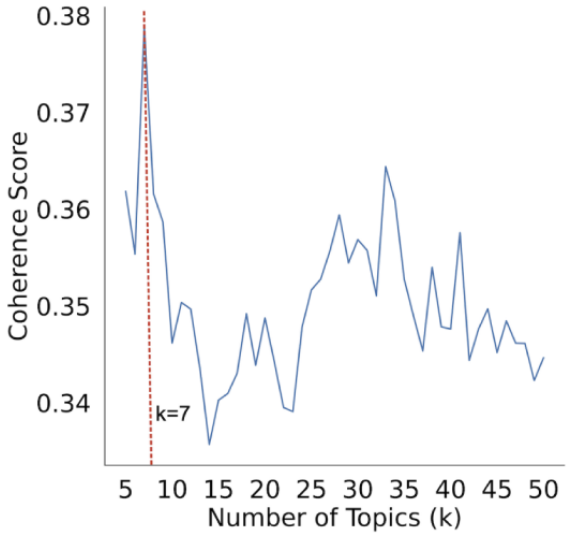
Topic modeling techniques have been adopted to analyze and identify key themes within social media discourse [29-32] including AD/ADRD-related discussions [22,23]. We built topic models on our entire dataset (posts and comments) by using latent Dirichlet allocation (LDA) [33], an unsupervised machine learning algorithm widely used for analyzing corpora. LDA generates latent topic distributions, making it a valuable tool for analyzing large, unlabeled document collections and their clustering into distinct groups based on common themes. LDA has been heavily leveraged in previous work on social media language analysis [29,30]. After tokenization, stop word removal, and adding n-grams (n=2,3), we created bag-of-words representations of the dataset.

Determining the Optimal Number of Topics

Given that LDA does not automatically determine the optimal number of topics, we adopted a semiautomated approach, motivated by previous work [22,30,31]. For this, we first used the coherence score as a metric to assess the model fit, varying the number of topics. The coherence score evaluates the tendency of words within a topic to co-occur and correlates well with topic modeling quality [34]. Then, we hand-selected a few topic models that occurred in the zone of highest coherence scores. Two coauthors manually evaluated the quality of these different topic models in terms of assessing the between-topic heterogeneity and within-topic homogeneity of words to narrow down the best topic model for our ensuing analyses.

We computed the coherence score for our 3 preprocessed corpora, varying the number of topics (k) from 5 to 50 (refer to Figure 1). We found the highest coherence at k=7. Consequently, we manually evaluated the quality of the topic modeling for k=5, k=7, and k=10.

Figure 1. Coherence values over varying number of topics in the LDA (latent Dirichlet allocation) topic model.



Labeling the Topics

The above steps yielded topical clusters, each of which contained clusters of keywords. We adopted a thematic analysis approach to label each topical cluster in a meaningful and interpretable way. The thematic analysis was conducted by 2 authors, familiar with social media analysis and AD/ADRD literature. They adopted an inductive coding approach to first independently code the topics based on the cluster of keywords, referencing sample posts corresponding to each topic. Then,

the authors engaged in collaborative discussions, comparing their codes, and gradually converging on cohesive thematic labels. This meticulous process ensured a nuanced and comprehensive analysis of the data. We found that the topic modeling with k=7 yielded the best results in terms of coherency within topics and boundaries between topics.

Classifying Support and Examining Supportive Expressions

Toward answering RQ2, we analyzed the responses within the online communities to examine the nature and forms of social support provided. This section details our methodology for categorizing these responses into distinct subcategories and contrasts the patterns of support provision within the Alzheimer community against those observed in a broader user demographic.

Classifying Support Expressions

Overview

The role of social support is essential in assisting individuals in managing psychological distress and various life adversities [35]. The emergence of social media platforms and virtual communities has led to the digital transposition of traditional social support mechanisms [13,36]. Specifically, online forums such as Reddit have increasingly mirrored the characteristics of virtual support groups, providing a platform for communal interaction and aid. The SSBC schema [37] offers a structured approach for classifying various types of support. Within this framework, informational support (IS) and emotional support (ES) emerge as 2 pivotal categories. Notably, these forms of support have garnered significant theoretical and empirical interest within social computing, as documented in various studies [20,38-43]. ES encompasses actions that provide encouragement, empathy, or care, whereas IS consists of assistance involving advice, information, or knowledge [37]. These 2 forms of support are particularly dominant in online interactions, as they can be effectively conveyed through text, making them well-suited for digital environments where individuals seek both guidance and emotional reassurance.

We categorized support offered in comments into ES and IS across both Reddit and ALZConnected datasets. To achieve this, we used a transfer learning methodology, a process wherein a machine learning classifier is developed by transferring insights from one labeled dataset to another unlabeled dataset with similar characteristics [44]. Specifically, we used a labeled Reddit dataset, previously expert-annotated with ES and IS labels on 396 comments collected from 55 mental health-related subreddits on Reddit [40]. This served as the seed dataset to train our classifiers. In addition, we manually labeled the ES and IS of 100 comments from AlzConnected and 100 comments from r/Alzheimers datasets and added these to the training datasets of the support classifiers. Our initial classifier was a supervised learning classifier for ES and IS, using n-grams (with n=1,2) as features in alignment with methodologies outlined in previous research [39-41]. We adopted a k-fold cross-validation (k=10) approach for evaluating our models.

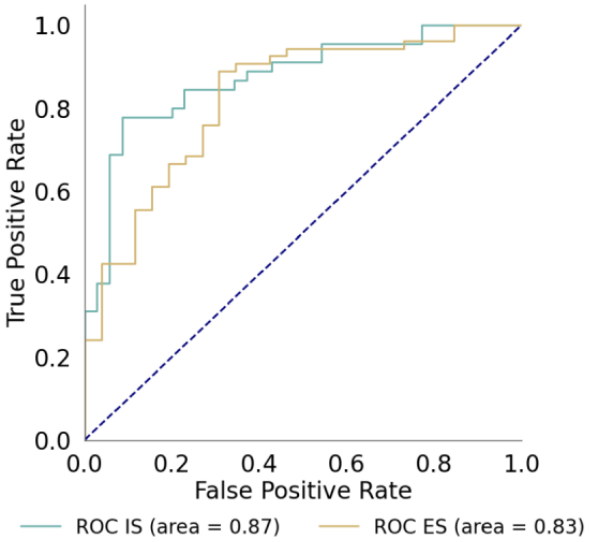
We used a range of classification algorithms, including naive Bayes, logistic regression, support vector machine (SVM), random forest, and neural network. Among these, the Linear SVM model demonstrated the best performance, achieving a mean area under the curve of 0.84 for ES and 0.87 for IS (refer to Table 2). Figure 2 provides the receiver-operating-characteristic area-under-curve plot for the classifiers used in identifying support types. We use the linear SVM model for our ensuing analyses. For the ES classifier, the predominant features include: “good,” “sorry,” “right,” “dont,” “going,” “people,” “get,” “life,” “im,” and “feel.” In contrast, the IS classifier highlighted features such as: “help,” “might,” “thing,” “go,” “think,” “try,” “get,” “see,” “probably,” and “doesnt.”

Table 2. Performance metrics of machine learning models of social support classifiers.

Model	Emotional support			Informational support		
	Precision	Recall	Area under the curve	Precision	Recall	Area under the curve
Naive Bayes	0.70	0.58	0.80	0.67	0.66	0.76
Logistic regression	0.73	0.69	0.76	0.75	0.76	0.81
Support vector machine	0.74	0.74	0.80	0.73	0.73	0.83
Random forest	0.68	0.54	0.75	0.67	0.66	0.75
Neural network	0.71	0.68	0.79	0.70	0.65	0.77
After adding augmented labeled data from AD/ADRD[a] communities						
Logistic regression	0.74	0.72	0.77	0.74	0.76	0.80
Support vector machine	0.75	0.76	0.83	0.77	0.78	0.87

[a]AD/ADRD: Alzheimer disease and related dementias.

Figure 2. Receiver-operating-characteristic area-under-curve plot of the emotional and informational support classifiers. ES: emotional support; IS: informational support.



To validate the applicability and relevance of the model, we applied transfer-based learning to the classifiers within our dataset. We conducted a manual validation on a randomly selected subset of 100 comments. This process involved 2 independent annotators classifying each reply for ES and IS. Discrepancies in annotations were resolved through the intervention of a third adjudicator. Further, we compared these manual annotations with the predictions made by the machine learning classifiers. The comparison resulted in an F_1-score of 0.76 for ES, characterized by a precision of 0.74 and accuracy of 0.76, and an F_1-score of 0.77 for IS, precision of 0.77, and accuracy of 0.77, which is an increase from the initial classifier. Given these results, we consider the classifier's performance satisfactory for further analysis. It is important to note that in our classification, ES and IS are not mutually exclusive categories; a single comment may be either both, or neither forms of support. Table 3 shows some example comments in our dataset with labels of ES and IS.

Table 3. Example comments and support labels with emotional and informational support (each comment may be either both, or neither forms of support).

Example comment	Emotional support	Informational support
Just as a note, when my Uncle passed away, I still remember him as the strong figure versus any thoughts of his Alzheimer. Man, my condolences and it's more of a blessing for him and your family for the rapid decline.	✓	x
Lumbar punctures are great and the data is very solid in the ability to diagnose AD[a]. However, LPs can be lengthy (and therefore expensive) procedures and are generally viewed as invasive by the general population. A blood test such as this would highly impact our ability to screen people more quickly for drug trials :)	x	✓
Sometimes you just have to chuckle. I just received yet another call from my person with dementia telling me they needed someone to fix their phone. The same phone they made a call with. And we live over three hours away.	✓	✓
This feels like a lecture, which is not what I need right now. No reply is necessary.	x	x

[a]AD: Alzheimer disease.

Associating the Relationship Between Topics of Concern and Support Received

To assess the support associated with each topic, we used linear regression models to predict the likelihood of receiving different types of support. We built separate models for ES and IS. Each model was designed to predict the probability of receiving a specific type of support based on the topic being discussed. So, we used the topics in the posts' content as independent variables, and the received support as the dependent variable. Essentially, by examining ordinary least squares regression models, we aimed to understand how different topics are associated with the probability of receiving a type of support. This approach allowed us to quantify the relationship between topics and the support they generate, providing valuable insights into the dynamics of support within the community.

Ethical Considerations

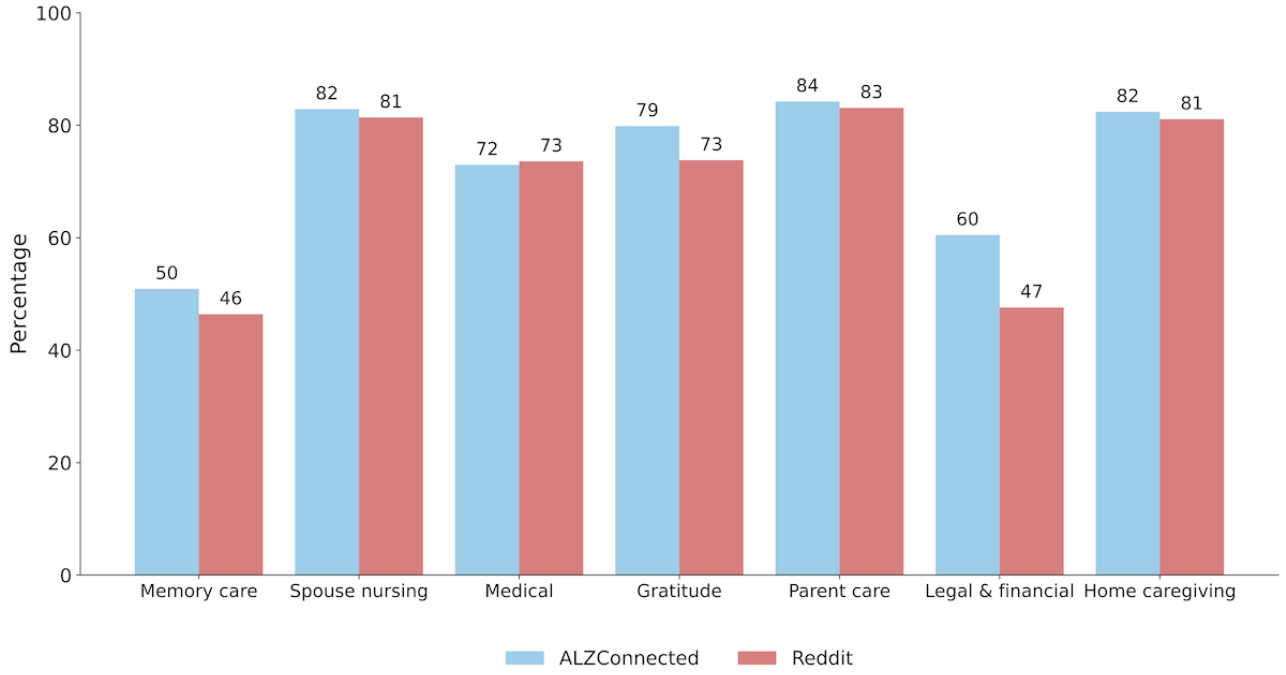
This project used publicly available data from online communities, and did not involve any direct interaction with human subject participants. Therefore, this research did not require ethics board approval. However, we are committed to the ethics of the research and followed practices to secure the privacy of the individuals in our dataset. We recognize the sensitivities of our study in terms of revealing the identities of the individuals, and thus deidentified all collected data. This paper only presents paraphrased quotes to reduce traceability while providing necessary context.

Results

RQ1: Themes of Needs and Concerns in Online Communities

We now describe the findings from our topic modeling followed by thematic analyses. Table 4 shows the 7 topics and relevant thematic analysis derived through LDA topic modeling. These topics serve as the primary themes, summarizing the needs and concerns of members within the AD/ADRD online communities. Figure 3 shows the overall distribution of topical occurrences in the 2 datasets, showing that the topics appear similarly on both platforms.

Figure 3. Distribution of topics as a percentage of posts on r/Alzheimers and ALZConnected.



We describe these topical themes as: First, the topics reveal diversity in community members' needs based on relationships with patients experiencing AD/ADRD. We find that several topics, such as memory care support, nursing for a spouse, caring for patients, and caregiving at home are linked to informal caregivers, such as patients' spouses, children, and close family members, along with professional caregivers providing in-home assistance. This reveals that community members with different roles may encounter various concerns and needs.

We find a substantial amount of content focusing on memory care, particularly about early onset patients with AD/ADRD, with words like "memory care" and "assisted living facility" (Topic 1, 5, and 7). In general, memory loss is a common symptom of AD/ADRD [45]. Memory care is frequently observed as dedicated and free-standing memory care living spaces within assisted living communities or integrated into life care communities [46]. Individuals also discuss how to choose an assisted living home with a memory care unit for patients and issues encountered in memory care units. One individual posted in r/Alzheimers to seek suggestions:

> *Hello everyone, could we discuss the options for finding an assisted living facility with memory care? My wife has Early-onset Alzheimer, and she is 57.*

We also find members who are early-stage patients looking for advice to plan the future and are gaining insight into experience within the memory care unit. An early onset patient posted:

> *Next week, I plan to visit an assisted living home with a memory care unit. My doctor has advised me to begin preparing for my future no matter what I can.*

AD/ADRD leads to a progressive decline in the capacity to perform daily living activities [47]. We found extensive discussions related to nursing and care, with the majority originating from caregivers and patients' family members (Topic 2, 5, and 7). Some content involves sharing AD/ADRD-related books and information. Within these posts, members recommend or summarize useful knowledge in a comprehensive way, and they are enthusiastic about discussing it with others. For instance, one caregiver posted their thoughts on the recommended book, "I recommend giving this book a read. Feel free to share your thoughts with me if you explore it." Also, a significant amount of content revolves around the sharing of experiences and challenges. Some members shared experiences when they felt happy in a largely challenging journey of caregiving for their family members with Alzheimer. For instance, a member wrote;

> *He didn't know who I was, but he got very excited that I helped him make handprint art to take to his 2 daughters and 3 granddaughters. He may not recognize me, but things like this show me the love is still there. Feeling blessed.*

Table 4. Topic themes with representative keywords and example posts.

	Topic keywords	Example post
Topic 1: Memory care and assisted living	care, facility, memory, memory care, wife, home, living, darling, dear darling, dear, darling wife, dear darling wife, assisted, assisted living, nursing, staff, significant, living facility, assisted living facility, and room	*Is there anyone willing to share their experience of placing their loved one into a memory care unit?*
Topic 2: Nursing for spouse	significant, director, director nursing, nursing, just, husband, know, dear, time, like, day, good, dear husband, think, sorry, hope, things, did, going, and got	*If I get COVID, would my wife have to be placed in a nursing home? What actions should I do in such a situation?*
Topic 3: Medical condition and diagnosis	years, alzheimer, disease, husband, ago, alzheimer disease, significant, diagnosed, year, just, stage, diagnosis, early, time, years agodementia, months, know, care, and memory	*What support or resources are available for people in the MCI stage who may be unaware of it? I'm asking because my younger sister has been diagnosed with early-onset Alzheimer.*
Topic 4: Acknowledgment and gratitude	significant, thank, dementia, good, hospice, help, loved, support, meds, know, doctor, thanks, best, sorry, disease, helpful, forum, new, and information	*I'd like to thank those who recommended getting an electronic pet for a loved one with Alzheimer. I bought one for my grandma. She absolutely loves it!*
Topic 5: Caring for parents	mom, just, significant, like, time, help, know, mother, things, home, going, day, really, think, doesn, house, need, feel, dad, and tell	*I am an only daughter and I have taken on the responsibility of caring for my mother who received a diagnosis of Alzheimer and vascular dementia for 2 years.*
Topic 6: Legal and financial matters	attorney, husband, dear husband, care, dear, power, power of attorney, need, law, Medicaid, significant, help, elder law, money, long, state, law attorney, elder law attorney, and pay	*My husband was just diagnosed with younger onset Alzheimer. We are in search of an elder care lawyer in the Naperville or surrounding area to assist us with financial planning.*
Topic 7: Caregiving at home	dementia, dad, person, mom, family, care, make, person dementia, people, like, time, life, significant, need, help, things, think, caregiver, nursing, and home	*I have taken on the role of a "mom" for my mom; she relies on me for all her necessities, and needs, or just for chat. I'm only 20. I wish to live my own life. However, guilt keeps me at home, and I'm concerned that moving out might intensify it. Please tell me what to do.*

Some members place individuals with AD/ADRD in skilled nursing homes, while others serve as caregivers at home, leading to varying sets of issues. Some common concerns in nursing homes are the occurrence of living issues faced by patients residing in facilities, inappropriate interactions among patients, and instances where patients get expelled from the nursing facility and sent back home. We also found an instance where the patient required hearing aids, but the care provided by nursing facilities was not satisfactory. Consequently, a member posted:

> *Regrettably, there are many instances where the thing is either not carried out properly or not done at all. During my visits, I've observed that there are times when she has the hearing aids in, but they are non-functional.*

In-home caregivers share the fatigue and strain arising from family relationships, as well as the distress of balancing their own lives with caregiving responsibilities. One member posted about being the sole caregiver without other siblings' assistance. They wrote, seeking advice:

> *I have a full-time job and my own family. I find myself in need of a break, and I'm hesitant to consider placing her in a nursing home. I'm concerned that the unfamiliar environment might exacerbate her dementia.*

Community members also expressed gratitude, encouragement, and acknowledgment. Members within online AD/ADRD communities can communicate about specific health-related topics, often finding online support a welcoming and comfortable venue for discussing sensitive issues [48]. Part of this involves mutual support and appreciation among community members, often occurring when someone shares their story or receives help and support from other members. Notably, posts and comments of this nature are typically longer and more detailed, prompting many individuals to express gratitude at the end, thanking others for taking the time to read their detailed posts. For instance, a member on r/Alzheimers, after venting about the challenges of being a caregiver, expressed:

> *Thanks for the opportunity to share my thoughts right here. I'll read your many posts before mine and offer my opinions and encouragement.*

Individuals also express gratitude not only directed toward other community members but also between caregivers and patients. Following the death of a loved one, members grieved recalling their shared past, the challenges of life, and caregiving after the diagnosis, such as, "I am thankful that his suffering finally ends. He is with my mom now in heaven." Furthermore, when reminiscing about early shared experiences, members often express gratitude and love for the individuals with AD/ADRD, such as, "Reflecting on the cherished moments we shared brings me joy, and I am grateful for the knowledge he imparted to me."

Finally, we also observed concerns related to legal and financial matters. Many of these concerns come from family members of individuals with AD/ADRD. With the progression of AD/ADRD, patients require increasing levels of care and assistance, leading to legal issues such as establishing guardianship or power of attorney (POA). Besides, financial concerns are a widely discussed topic in the communities, covering aspects such as handling patient's assets, as well as the costs associated with caregiving and treatment. For example, a member posted to seek advice from the community:

> *Her funds is running down. We still hope that we can place her in a nursing facility after her funds are used up by the memory care (MC) facility. I need to know if there is any assistance available to support her continued stay in MC.*

RQ2: Social Support in Online AD/ADRD Communities

Distribution of ES and IS

We first examined how ES and IS are distributed in our datasets. We observe similar trends in the distribution of support in both communities—on average, IS is 6% more prevalent in r/Alzheimers and 10% more prevalent in AlzConnected than ES in respective communities. Figure 4 shows the overall distribution of ES and IS in the 2 datasets.

As previously noted, we obtained the topics on the posts and the support received in the comments to the posts. We break this down into the topics of the posts, and Figure 5 shows the distribution of ES and IS in the comments to the posts of varying topics in the 2 datasets.

Both Figures 4 and 5 also show independent sample *t* tests between the occurrences of ES and IS. For each topical theme, we note statistically significant differences in occurrences of the 2 types of support.

Figure 4. Distribution of ES and IS to posts on r/Alzheimers and ALZConnected.

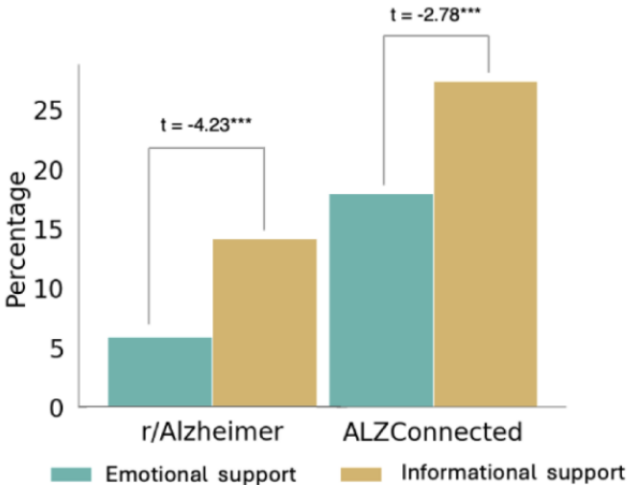
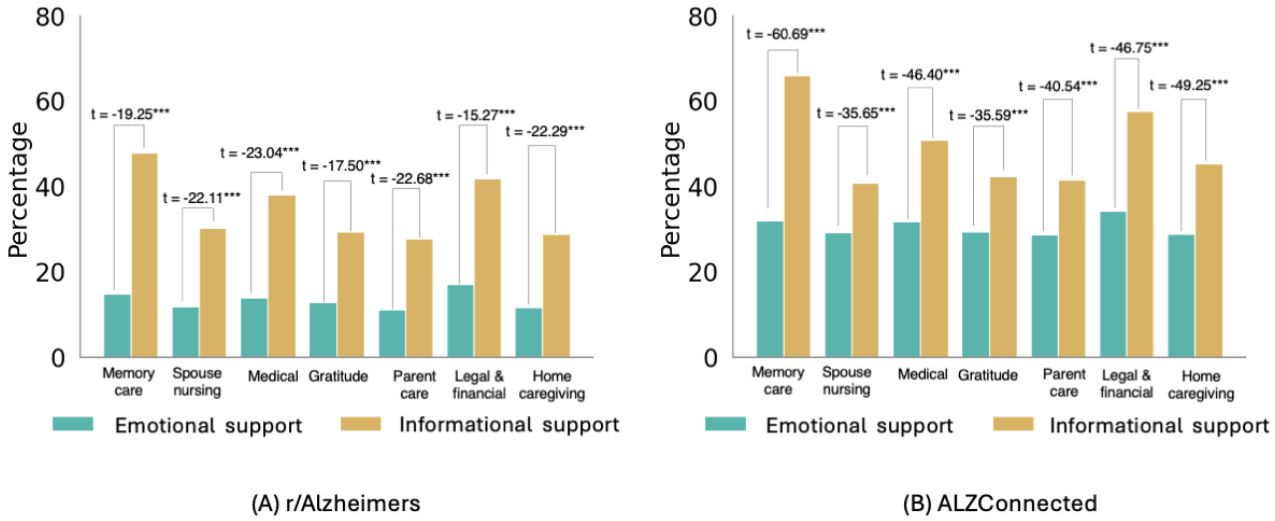


Figure 5. Distribution of emotional and IS in comments to posts with varying topics on (A) r/Alzheimers and (B) ALZConnected.



Regression Models for Predicting Support

To further analyze how support was provided corresponding to each topic, we built regression models. For each of our datasets (r/Alzheimers and AlzConnected), we built 2 models, 1 with IS as the dependent variable and another with ES as the dependent variable. Both of the models had all the topic occurrences as independent variables. Essentially, these models examined the likelihood of receiving ES and IS for each topic.

Table 5 shows the coefficients of the independent variables (topics) for ES and IS across the 2 datasets of r/Alzheimers and AlzConnected. All of these models show a significant goodness-of-fit (R^2). These coefficients signal the likelihood of receiving the respective support in the community. We describe our observations below.

Table 5. Coefficients of linear regression models with the topics of posts as independent variables and support (emotional and informational) responses as dependent variables.

Independent variables	Dependent variables, coefficient			
	Emotional support		Informational support	
	r/Alzheimers	ALZConnected	r/Alzheimers	ALZConnected
Intercept	–3.00E-04	3.00E-04	–0.008[a]	0.027[a]
Memory care	0.014	0.012[b]	0.191[a]	0.258[a]
Nursing for spouse	0.016[b]	0.079[a]	0.066[a]	0.013[a]
Medical condition	0.048[a]	0.067[a]	0.182[a]	0.164[a]
Acknowledgment and gratitude	0.041[a]	0.058[a]	0.038[a]	0.034[a]
Caring for parents	0.014[c]	0.089[a]	0.031[a]	0.048[a]
Legal and financial	0.056[a]	0.075[a]	0.088[a]	0.13[a]
Caregiving at home	0.035[a]	0.038[a]	0.055[a]	0.129[a]
R^2	0.777[a]	0.144[a]	0.260[a]	0.294[a]

[a]$P<.001$.
[b]$P<.01$.
[c]$P<.05$.

ES Posts and Topics

For both r/Alzheimers and AlzConnected, we find that topics relating to nursing and caring, such as “Nursing for Spouse,” “Caring for Parents,” and “Caregiving at Home” show high positive coefficients with statistical significance, that is, posts with these topics are likely to receive ES. Further, topics that are seemingly about informational content, such as “Legal and Financial” and “Medical Conditions and Diagnosis” also receive ES.

For instance, an individual self-disclosed their stress about the fact that their mom changed a POA and was in complete denial of her symptoms:

> *I believe my mom just changed her POA and took me completely off. Her husband is covering for her and is in complete denial about her diagnosis. [..] She is clearly paranoid and delusional. I will consult with an elder care attorney soon. Any other advice? She blames me for her driver's license getting taken away. I am the only family member in the state. Neither she nor her husband are capable of overseeing her medical care. [..] I am very stressed about this situation. Thank you for any thoughts.*

IS Posts and Topics

For both r/Alzheimers and AlzConnected, posts relating to "Memory Care Support," "Medical Condition and Diagnosis," "Legal and Financial," and "Caregiving at Home" show high positive coefficients with statistical significance. Interestingly, posts relating to "Nursing for Spouse" and "Caring for Parents," which occurred high for ES, are also likely to receive IS. For instance, an individual who described their situation, also asked, "How are folks keeping their careers while helping to care for a parent?"

Discussion

Principal Results

This study sheds light on the complex needs of individuals dealing with AD/ADRD as patients, caregivers, family members, and loved ones. Our analyses revealed seven primary topical themes—(1) memory care and assisted living, (2) nursing for a spouse, (3) medical condition and diagnosis, (4) acknowledgment and gratitude, (5) caring for parents, (6) legal and financial aspects, and (7) caregiving at home. We found that IS was more prevalently provided than ES across both platforms. In particular, ES was predominantly found in posts related to nursing and caregiving, suggesting that caregivers turn to these communities to cope with the emotional burdens associated with their roles. In addition, IS was more frequently associated with concerns about medical conditions, legal and financial issues, and caregiving at home, indicating a demand for practical advice and resources in managing the complexities of AD/ADRD.

These findings highlight the critical role that online communities play in supplementing the support networks of AD/ADRD caregivers and family members. Many caregivers face challenges in accessing adequate in-person support due to stigma, social isolation, or lack of resources, making online platforms an important alternative for both emotional validation and practical advice—as also noted in the case of other stigmatized mental health [13,38-41,49] and health conditions (such as cancer [17,19,20] and HIV [50]). The predominance of IS suggests that individuals seek concrete advice to manage the daily realities of caregiving, including navigating health care systems, legal documentation, and financial planning. This aligns with previous research indicating that caregivers often experience a steep learning curve when managing AD/ADRD, requiring access to timely and accurate information [51,52]. Furthermore, the need for ES around nursing and caregiving emphasizes the significant emotional and psychological toll on caregivers [53]. These findings underscore the importance of community support in mitigating mental health issues such as feelings of burnout, stress, and isolation—some of the major concerns experienced by AD/ADRD caregivers [54]. We also noted expressions of gratitude and acknowledgment within these communities, which may indicate a sense of belonging and solidarity among the members of online communities, especially among caregivers to help boost their resilience and sustain their caregiving roles over the long term. Therefore, fostering these online communities can be a low-cost and scalable way to address caregiver burnout, which is a critical issue given the growing aging population and the increasing prevalence of AD/ADRD.

Our study also bears important implications for health care practitioners, policy makers, and support organizations. First, online communities provide empirical insights into the plausible gaps in accessible resources for caregivers. For instance, institutional support from governmental policies, health care systems, and AD/ADRD care organizations can leverage online platforms to identify the needs. Further, these platforms can be used to disseminate accurate and specific information tailored to the needs of the caregivers, potentially reducing the burden on healthcare professionals and improving caregiver confidence in managing AD/ADRD-related challenges. In addition, these platforms can be integrated into existing health care support structures and information.

Health care systems and AD care organizations could leverage online platforms to disseminate accurate, easily digestible information tailored to the needs of caregivers, potentially reducing the burden on health care professionals and improving caregiver confidence in managing AD/ADRD-related challenges. In addition, these platforms could be integrated into existing health care support structures, allowing caregivers to receive timely advice from health care professionals, thus improving patient outcomes. From a policy perspective, there is a need to enhance public health initiatives that support caregivers, especially in terms of financial planning and legal guidance, which was one of the major topics of discussion.

Policy makers could consider developing targeted programs that address these specific needs, such as offering workshops on financial planning for families affected by AD/ADRD or increasing access to legal assistance. Moreover, expanding caregiver support programs, including respite care and mental health services, could alleviate the emotional burden as prevalently discussed in the online communities. For technology designers and developers, our study suggests opportunities to enhance digital tools aimed at supporting caregivers. In particular, with the increasing prevalence of using artificial intelligence (AI) in different sectors, platforms can also incorporate AI-driven recommendation systems to provide personalized resources based on each individual's specific question or concern. In addition, integrating mental health support tools, such as access to counseling services or peer support groups within these online communities, could further enhance the mental well-being of caregivers.

Limitations

Our study has limitations which also suggest interesting future research directions. Although adopting computational

approaches enabled our exploration of the depth and breadth of support dynamics across various topics, our findings are not necessarily generalizable to the broader population or other contexts beyond the online communities analyzed. As noted in Figure 3, although both communities follow the trend of having more IS posts than ES, the overall level of engagement across the 2 communities is not the same. In particular, our study is inherently limited by self-selection bias, as we only study individuals who actively chose to post and engage in these online communities. This means that our findings may disproportionately reflect the experiences of individuals who are more active on social media and are motivated and willing to share publicly on these platforms. In addition, this study does not capture the perspectives of passive users who browse forums for information without directly participating, which may exclude a significant portion of the online communities. Furthermore, digital inequity presents another challenge. Not all individuals have equal access to digital platforms, reliable internet, or the necessary skills to engage effectively online. This disparity can exclude voices from socioeconomically disadvantaged or digitally marginalized groups, leading to a biased understanding of support dynamics in online communities.

Accordingly, future research should explore strategies to mitigate these limitations by examining the representativeness of the user base and incorporating voices from underrepresented groups. For example, comparing findings from active participants with those of passive users or supplementing online data with offline sources could provide a more comprehensive view of support mechanisms. In addition, understanding how users of online support communities differ from those who might use AI-driven tools for AD/ADRD management is critical. Exploring the overlap and divergence between these groups will help inform the design of inclusive and effective digital interventions. Finally, the emergence of generative AI and large language models has enabled the development of personalized assistants and chatbots [55,56]. It would be interesting to examine how these AI chatbots respond to people's queries on AD/ADRD and how effective they are compared to online communities.

Comparison With Previous Work

Previous research has highlighted the significant burden imposed by AD/ADRD on patients, caregivers, and health care systems [54,57,58]. These caregivers often face stigma due to societal misconceptions about caregiving, which can result in judgment and disapproval, further intensifying their feelings of isolation and stress [59,60]. The growing demand for caregiving support due to the neurodegenerative progression of AD/ADRD necessitates increased levels of care over time that by 2030, an additional 1.2 million direct care workers will be required in the United States to address the rising needs of this population, pointing to a substantial gap in the current AD/ADRD caregiving infrastructure [2]. Our findings offer empirical evidence on the specific needs and concerns of individuals affected by AD/ADRD as they seek and exchange support within online communities.

Prior research has also underscored the psychosocial burdens faced by AD/ADRD caregivers, including feelings of marginalization and neglect of well-being [60,61]. A total of 69% of caregivers experience moderate to severe levels of caregiving burden, with the severity of the patient's condition significantly impacting their stress levels [62]. Our study expands on this by underlining the specific concerns expressed in online communities and revealing that discussions frequently center on memory care, nursing, caregiving challenges, and legal and financial issues. This suggests that caregivers seek both practical advice as well as ES to manage the complexities of their roles.

A notable contribution of our research is its focus on the role of online platforms as a source of support for AD/ADRD caregivers. Previous studies have acknowledged the use of digital tools among caregivers [63] and highlighted a disconnect between the availability of digital support and the recognition of caregiving efforts within social circles [64]. Our findings on the major concerns and needs—as identified by topic modeling—also align with recent computational and qualitative analyses of Reddit and other online communities for AD/ADRD caregivers, which have identified similar themes, including ES-seeking, decision-making about care facilities, and management of behavioral symptoms [23,27]. Therefore, the integration of individuals with AD/ADRD and caregivers' needs into the development of health and wellbeing technologies is necessary for future design [9,64]. Our study contributes to the recent body of work on the role of technologies in supporting caregivers—Bhat et al [65] emphasized the crucial role of caregiver-focused technological supports, and Kim et al [66] highlighted the need for tailored support strategies caregivers need during different phases of challenging behavioral episodes, and Meyerhoff et al [67] stressed the importance of user-centered digital mental health tools that adapt to individual needs.

The high prevalence of IS identified in our study aligns with existing literature that suggests caregivers often turn to digital platforms to fill gaps in health care services, particularly due to inconsistent coordination of long-term care [4]. The distribution of ES and IS across topics highlights a nuanced understanding of how caregivers use online communities to address specific facets of the caregiving experience. Aligning with how previous work focused on the use of technology for caregiving support [3,4], our study revealed the unique dynamics of social media as a tool for fostering a sense of belonging, solidarity, and social support among AD/ADRD caregivers. These observations align with findings from studies by Rains et al [68] and Naslund et al [69], who noted that individuals often leverage online spaces to combat loneliness and build supportive relationships. Together, our study builds upon the foundation laid by the SSBC, which categorizes support into 5 types—informational, emotional, esteem, tangible, and social network support [24]. Although this study focused specifically on ES and IS, we observed overlaps with other forms of support as well.

Conclusions

This study reveals that online communities centered on AD/ADRD engage in extensive discussions on a diverse array of topics. These topics include memory care strategies, nursing and caregiving practices, as well as legal, financial, and emotional challenges associated with managing the disease. The findings highlight not only the primary concerns and pain points experienced by caregivers and families who manage AD/ADRD in their homes but also the critical role online platforms play in providing them with a sense of community, practical advice, and ES. By analyzing these discussions, the study identifies how individuals affected by AD/ADRD turn to online communities to share personal experiences, seek solutions to caregiving dilemmas, and obtain validation and reassurance from others who are navigating similar challenges. These online interactions can alleviate feelings of isolation, foster a collaborative environment for exchanging best practices, and serve as an essential resource for information and guidance. Our findings underscore the urgent need for targeted institutional and social interventions that address the nuanced and evolving needs of AD/ADRD patients, their caregivers, and their families. These interventions could include enhanced support services, caregiver education programs, and accessible mental health resources that are tailored to the specific challenges highlighted in these online discussions. By recognizing and addressing these needs, stakeholders can improve the quality of life for both patients and caregivers, ultimately contributing to a more supportive and inclusive environment for those affected by AD/ADRD.

Conflicts of Interest

None declared.

References

1. Montenegro JMF. Alzheimer's disease diagnosis based on cognitive methods in virtual environments and emotions analysis. arXiv:1810.10941 2018.
2. NA. 2023 alzheimer's disease facts and figures. Alzheimers Dement 2023;19(4):1598-1695. [doi: 10.1002/alz.13016] [Medline: 36918389]
3. Rosebush CE, Stabler H, Nkimbeng M, Louwagie K, Fields NL, Jutkowitz E, et al. The porchlight project: A pilot study to adapt the senior companion program to enhance memory care services and supports. Gerontol Geriatr Med 2021;7:23337214211017651 [FREE Full text] [doi: 10.1177/23337214211017651] [Medline: 34036120]
4. Borson S, Chodosh J. Developing dementia-capable health care systems: a 12-step program. Clin Geriatr Med 2014;30(3):395-420. [doi: 10.1016/j.cger.2014.05.001] [Medline: 25037288]
5. Farhana N, Peckham A, Marani H, Roerig M, Marchildon G. The social construction of dementia: implications for healthcare experiences of caregivers and people living with dementia. J Patient Exp 2023;10:23743735231211066 [FREE Full text] [doi: 10.1177/23743735231211066] [Medline: 38026064]
6. Klimova B, Maresova P, Valis M, Hort J, Kuca K. Alzheimer's disease and language impairments: social intervention and medical treatment. Clin Interv Aging 2015;10:1401-1407 [FREE Full text] [doi: 10.2147/CIA.S89714] [Medline: 26346123]
7. El Haj M, Jardri R, Larøi F, Antoine P. Hallucinations, loneliness, and social isolation in alzheimer's disease. Cogn Neuropsychiatry 2016;21(1):1-13. [doi: 10.1080/13546805.2015.1121139] [Medline: 26740416]
8. Sun Y, Ji M, Leng M, Li X, Zhang X, Wang Z. Comparative efficacy of 11 non-pharmacological interventions on depression, anxiety, quality of life, and caregiver burden for informal caregivers of people with dementia: A systematic review and network meta-analysis. Int J Nurs Stud 2022;129:104204. [doi: 10.1016/j.ijnurstu.2022.104204] [Medline: 35247788]
9. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
10. Berridge C, Turner NR, Liu L, Karras SW, Chen A, Fredriksen-Goldsen K, et al. Advance planning for technology use in dementia care: development, design, and feasibility of a novel self-administered decision-making tool. JMIR Aging 2022;5(3):e39335 [FREE Full text] [doi: 10.2196/39335] [Medline: 35896014]
11. Rettinger L, Zeuner L, Werner K, Ritschl V, Mosor E, Stamm T, et al. A mixed-methods evaluation of a supporting app for informal caregivers of people with dementia. 2020 Presented at: PETRA '20: The 13th PErvasive Technologies Related to Assistive Environments Conference; 2020 June 30-3 July; Corfu Greece p. 1-8. [doi: 10.1145/3389189.3397981]
12. Queluz FNFR, Kervin E, Wozney L, Fancey P, McGrath PJ, Keefe J. Understanding the needs of caregivers of persons with dementia: a scoping review. Int Psychogeriatr 2020;32(1):35-52 [FREE Full text] [doi: 10.1017/S1041610219000243] [Medline: 30967164]
13. De Choudhury M, De S. Mental health discourse on reddit: self-disclosure, social support, and anonymity. ICWSM 2014;8(1):71-80. [doi: 10.1609/icwsm.v8i1.14526]
14. Ernala SK, Rizvi AF, Birnbaum ML, Kane JM, De Choudhury M. Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia. Proc. ACM Hum.-Comput. Interact 2017;1(CSCW):1-27. [doi: 10.1145/3134678]

15. Jin Y, Chandra M, Verma G, Hu Y, Choudhury MD, Kumar S. Better to ask in english: cross-lingual evaluation of large language models for healthcare queries. 2024 Presented at: WWW '24: Proceedings of the ACM Web Conference 2024; 2024 May 13 - 17; Singapore p. 2627-2638. [doi: 10.1145/3589334.3645643]
16. Yang D, Yao Z, Kraut R. Self-disclosure and channel difference in online health support groups. ICWSM 2017;11(1):704-707. [doi: 10.1609/icwsm.v11i1.14935]
17. Levonian Z, Erikson DR, Luo W, Narayanan S, Rubya S, Vachher P, et al. Bridging qualitative and quantitative methods for user modeling: tracing cancer patient behavior in an online health community. ICWSM 2020;14:405-416. [doi: 10.1609/icwsm.v14i1.7310]
18. Saha K, Ernala SK, Dutta S, Sharma E, De Choudhury M. Understanding moderation in online mental health communities. : Springer; 2020 Presented at: 12th International Conference, SCSM 2020, Held as Part of the 22nd HCI International Conference, HCII 2020; 2020 July 19–24; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49576-3_7]
19. Olteanu A, Varol A, Kiciman E. Distilling the outcomes of personal experiences: A propensity-scored analysis of social media. 2017 Presented at: CSCW '17: Computer Supported Cooperative Work and Social Computing; 2017 25 February; Portland Oregon USA p. 370-386. [doi: 10.1145/2998181.2998353]
20. Yang D, Yao Z, Seering J, Kraut R. The channel matters: self-disclosure, reciprocity and social spport in online cancer spport groups. 2019 Presented at: CHI '19: CHI Conference on Human Factors in Computing Systems; 2019 May 4 - 9; Glasgow Scotland Uk p. 1-15. [doi: 10.1145/3290605.3300261]
21. Lazar A, Brewer RN, Kacorri H, Hong J, Punzalan MND, Mahathir M, et al. How content authored by people with dementia affects attitudes towards dementia. Proc ACM Hum Comput Interact 2021;5 [FREE Full text] [doi: 10.1145/3479542] [Medline: 35187410]
22. Yoon S, Lucero R, Mittelman MS, Luchsinger JA, Bakken S. Mining twitter to inform the design of online interventions for hispanic alzheimer's disease and related dementias caregivers. Hisp Health Care Int 2020;18(3):138-143. [doi: 10.1177/1540415319882777] [Medline: 31646904]
23. Pickett AC, Valdez D, Sinclair KL, Kochell WJ, Fowler B, Werner NE. Social media discourse related to caregiving for older adults living with alzheimer disease and related dementias: computational and qualitative study. JMIR Aging 2024;7:e59294 [FREE Full text] [doi: 10.2196/59294] [Medline: 38896462]
24. Suhr JA, Cutrona CE, Krebs KK, Jensen SL. The social support behavior code (ssbc). In: Couple Observational Coding Systems. United Kingdom: Routledge; 2004:307-318.
25. Andalibi N, Haimson OL, Choudhury MD, Forte A. Social support, reciprocity, and anonymity in responses to sexual abuse disclosures on social media. ACM Trans. Comput.-Hum. Interact 2018;25(5):1-35. [doi: 10.1145/3234942]
26. Saha K, Kim SC, Reddy MD, Carter AJ, Sharma E, Haimson OL, et al. The language of LGBTQ+ minority stress experiences on social media. Proc ACM Hum Comput Interact 2019;3(CSCW):1-22 [FREE Full text] [doi: 10.1145/3361108] [Medline: 32935081]
27. Pickett AC, Valdez D, White LA, Loganathar P, Linden A, Boutilier JJ, et al. The careVirtue digital journal for family and friend caregivers of people living with alzheimer disease and related dementias: exploratory topic modeling and user engagement study. JMIR Aging 2024;7:e67992 [FREE Full text] [doi: 10.2196/67992] [Medline: 39719081]
28. r/alzheimersgroup. 2023. URL: https://knowyourmeme.com/memes/sites/ralzheimersgroup [accessed 2023-12-10]
29. Chancellor S, Lin Z, Goodman EL, Zerwas S, Choudhury MD. Quantifying and predicting mental illness severity in online pro-eating disorder communities. 2016 Presented at: CSCW '16: Computer Supported Cooperative Work and Social Computing; 2016 27 February-2 March; San Francisco California USA p. 1171-1184. [doi: 10.1145/2818048.2819973]
30. Zhou J, Saha K, Lopez Carron IM, Yoo DW, Deeter CR, De Choudhury M, et al. Veteran critical theory as a lens to understand veterans' needs and support on social media. Proc. ACM Hum.-Comput. Interact 2022;6(CSCW1):1-28. [doi: 10.1145/3512980]
31. Gong C, Saha K, Chancellor S. "The Smartest Decision for My Future": social media reveals challenges and stress during post-college life transition. Proc. ACM Hum.-Comput. Interact 2021;5(CSCW2):1-29. [doi: 10.1145/3476039]
32. Mehrotra R, Sanner S, Buntine W, Xie L. Improving LDA topic models for microblogs via tweet pooling and automatic labeling. 2013 Presented at: SIGIR '13: The 36th International ACM SIGIR conference on research and development in Information Retrieval; 2013 28 July-1 August; Dublin Ireland p. 889-892. [doi: 10.1145/2484028.2484166]
33. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. Cambridge, Massachusetts: MIT Press; 2003:993-1022.
34. Wallach H, Murray I, Salakhutdinov R, Mimno D. Evaluation methods for topic models. 2009 Presented at: ICML '09: The 26th Annual International Conference on Machine Learning held in conjunction with the 2007 International Conference on Inductive Logic Programming; 2009 June 14 - 18; Montreal Quebec Canada p. 1105-1112. [doi: 10.1145/1553374.1553515]
35. Kummervold PE, Gammon D, Bergvik S, Johnsen JK, Hasvold T, Rosenvinge JH. Social support in a wired world: use of online mental health forums in Norway. Nord J Psychiatry 2002;56(1):59-65. [doi: 10.1080/08039480252803945] [Medline: 11869468]
36. Rains S, Young V. A meta-analysis of research on formal computer-mediated support groups: examining group characteristics and health outcomes. Hum. Communication Research 2009;35(1):309-336. [doi: 10.1111/j.1468-2958.2009.01353.x]

37. Cutrona CE, Suhr JA. Controllability of stressful events and satisfaction with spouse support behaviors. Communication Research 1992;19(2):154-174. [doi: 10.1177/009365092019002002]
38. Andalibi N, Haimson OL, Choudhury MD, Forte A. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. 2016 Presented at: CHI'16: CHI Conference on Human Factors in Computing Systems; 2006 May 7 - 12; San Jose California USA p. 3908-3916. [doi: 10.1145/2858036.2858096]
39. De Choudhury M, Kiciman E. The language of social support in social media and its effect on suicidal ideation risk. ICWSM 2017;11(1):32-41. [doi: 10.1609/icwsm.v11i1.14891]
40. Sharma E, Choudhury MD. Mental health support and its relationship to linguistic accommodation in online communities. : ACM; 2018 Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; 2018 April 21 - 26; Montreal QC Canada p. 1-13. [doi: 10.1145/3173574.3174215]
41. Saha K, Sharma A. Causal factors of effective psychosocial outcomes in online mental health communities. ICWSM 2020;14:590-601. [doi: 10.1609/icwsm.v14i1.7326]
42. Saha K, Torous J, Caine ED, De Choudhury M. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. J Med Internet Res 2020;22(11):e22600 [FREE Full text] [doi: 10.2196/22600] [Medline: 33156805]
43. Wang Y, Kraut R, Levine JM. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. 2012 Presented at: CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work; 2012 February 11 - 15; Seattle Washington USA p. 833-842. [doi: 10.1145/2145204.2145329]
44. Torrey L, Shavlik J. Transfer learning. In: Handbook of research on machine learning applications and trends: algorithms, methods, and techniques. Hershey, Pennsylvania, USA: IGI global; 2010:242-264.
45. Jahn H. Memory loss in alzheimer's disease. Dialogues Clin Neurosci 2013;15(4):445-454 [FREE Full text] [doi: 10.31887/DCNS.2013.15.4/hjahn] [Medline: 24459411]
46. Calkins MP. Memory care and alzheimer's units. In: Environmental Psychology and Human Well-Being. Amsterdam, Netherlands: Elsevier; 2018:365-386.
47. Rolland Y, Pillard F, Klapouszczak A, Reynish E, Thomas D, Andrieu S, et al. Exercise program for nursing home residents with alzheimer's disease: a 1-year randomized, controlled trial. J Am Geriatr Soc 2007;55(2):158-165. [doi: 10.1111/j.1532-5415.2007.01035.x] [Medline: 17302650]
48. White M, Dorman SM. Receiving social support online: implications for health education. Health Educ Res 2001;16(6):693-707. [doi: 10.1093/her/16.6.693] [Medline: 11780708]
49. Moore D, Drey N, Ayers S. Use of online forums for perinatal mental illness, stigma, and disclosure: an exploratory model. JMIR Ment Health 2017;4(1):e6 [FREE Full text] [doi: 10.2196/mental.5926] [Medline: 28219879]
50. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a aystematic review of uses of social media in HIV communication. J Med Internet Res 2015;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
51. Soong A, Au ST, Kyaw BM, Theng YL, Tudor Car L. Information needs and information seeking behaviour of people with dementia and their non-professional caregivers: a scoping review. BMC Geriatr 2020;20(1):61 [FREE Full text] [doi: 10.1186/s12877-020-1454-y] [Medline: 32059648]
52. Kernisan LP, Sudore RL, Knight SJ. Information-seeking at a caregiving website: a qualitative analysis. J Med Internet Res 2010;12(3):e31 [FREE Full text] [doi: 10.2196/jmir.1548] [Medline: 20675292]
53. Blum K. Family caregivers. Palliat. care nursing. In: Qual. Care to End Life. Cham: Springer; 2010:187-198.
54. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Journal of Social Work Education 2017;44(sup3):105-113. [doi: 10.5175/jswe.2008.773247702]
55. Xu L, Sanders L, Li K, Chow JCL. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR Cancer 2021;7(4):e27850 [FREE Full text] [doi: 10.2196/27850] [Medline: 34847056]
56. Goodman S, Buehler E, Clary P, Coenen A, Donsbach A, Horne TN, et al. Lampost: design and evaluation of an ai-assisted email writing prototype for adults with dyslexia. 2022 Presented at: ASSETS '22: Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility; 2022 October 23 - 26; Athens Greece p. 1-8. [doi: 10.1145/3517428.3544819]
57. Gustavsson A, Norton N, Fast T, Frölich L, Georges J, Holzapfel D, et al. Global estimates on the number of persons across the alzheimer's disease continuum. Alzheimers Dement 2023;19(2):658-670. [doi: 10.1002/alz.12694] [Medline: 35652476]
58. Prince M. World Alzheimer report 2015. In: The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends. London: Alzheimer's Disease International; 2015.
59. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. Lancet Public Health 2022;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
60. Corrigan PW, Watson AC, Miller FE. Blame, shame, and contamination: the impact of mental illness and drug dependence stigma on family members. J Fam Psychol 2006;20(2):239-246. [doi: 10.1037/0893-3200.20.2.239] [Medline: 16756399]

61. Werner P, Goldstein D, Buchbinder E. Subjective experience of family stigma as reported by children of alzheimer's disease patients. Qual Health Res 2010;20(2):159-169. [doi: 10.1177/1049732309358330] [Medline: 20065304]
62. Pudelewicz A, Talarska D, Bączyk G. Burden of caregivers of patients with alzheimer's disease. Scand J Caring Sci 2019;33(2):336-341. [doi: 10.1111/scs.12626] [Medline: 30378698]
63. Siddiqui F, Varghese D, Singh P, Bayyavarapu SB, Lindsay S, Chandrasekara D, et al. Exploring the digital support needs of caregivers of people with serious mental illness. 2023 Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems; 2023 April 23 - 28; Hamburg Germany p. 1-16. [doi: 10.1145/3544548.3580674]
64. Chen Y, Ngo V, Park SY. Caring for caregivers: designing for integrality. 2013 Presented at: CSCW '13: Computer Supported Cooperative Work; February 23 - 27 2013; San Antonio Texas USA p. 91-102. [doi: 10.1145/2441776.2441789]
65. Bhat KS, Hall AK, Kuo T, Kumar N. "We are half-doctors": family caregivers as boundary actors in chronic disease management. Proc. ACM Hum.-Comput. Interact 2023;7(CSCW1):1-29. [doi: 10.1145/3579545]
66. Kim T, Kim H, Roberts AC, Jacobs M, Kay M. Opportunities in mental health support for informal dementia caregivers suffering from verbal agitation. Proc. ACM Hum.-Comput. Interact 2024;8(CSCW1):1-26. [doi: 10.1145/3637381]
67. Meyerhoff J, Kornfield R, Mohr DC, Reddy M. Meeting yung adults' social support needs across the health behavior change journey: implications for digital mental health tools. Proc ACM Hum Comput Interact 2022;6(CSCW2):1-33 [FREE Full text] [doi: 10.1145/3555203] [Medline: 36387059]
68. Rains SA. Coping with Illness Digitally. Cambridge, Massachusetts: MIT Press; 2018.
69. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS One 2014;9(10):e110171 [FREE Full text] [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]

Abbreviations

AD/ADRD: Alzheimer Disease and Related Dementias
AD: Alzheimer Disease
AI: artificial intelligence
ES: emotional support
IS: informational support
LDA: latent Dirichlet allocation
POA: power of attorney
RQ: research question
SSBC: social support behavioral code
SVM: support vector machine

Edited by A Garcia; submitted 18.11.24; peer-reviewed by AC Pickett; comments to author 11.02.25; revised version received 04.03.25; accepted 07.03.25; published 05.05.25.

Please cite as:
Kaliappan S, Liu C, Jain Y, Karkar R, Saha K
Online Communities as a Support System for Alzheimer Disease and Dementia Care: Large-Scale Exploratory Study
JMIR Aging 2025;8:e68890
URL: https://aging.jmir.org/2025/1/e68890
doi:10.2196/68890
PMID:

Determinants of Having Online Health Consultations During the COVID-19 Pandemic Among Middle-Aged and Older Adults in Germany: Representative Longitudinal Survey Study

Ariana Neumann, MSc; Hans-Helmut König, MPH, Prof Dr; André Hajek, Prof Dr

Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Martinistr. 52, Hamburg, Germany

Corresponding Author:
Ariana Neumann, MSc
Department of Health Economics and Health Services Research, University Medical Center Hamburg-Eppendorf, Martinistr. 52, Hamburg, Germany

Abstract

Background: During the COVID-19 pandemic, telemedicine services represented a widely implemented alternative to in-person doctor and therapist appointments. Consequently, rates of telemedicine use rapidly increased worldwide, also in Germany. Research regarding longitudinal determinants of telemedicine use is needed, particularly from nationally representative German samples, to improve understanding of the use behavior of major target groups such as middle-aged and older adults.

Objective: This study aimed to longitudinally investigate determinants of online health consultation use among middle-aged and older individuals during the COVID-19 pandemic in Germany.

Methods: Nationally representative longitudinal data of German middle-aged and older adults (≥46 years old) were taken from the German Ageing Survey (DEAS). Data from the Compact Survey (conducted between June and July 2020) and wave 7 (conducted between November 2020 and March 2021) of the DEAS were observed (pooled analytic sample N=5456). Having experienced consultations with doctors or therapists on online platforms served as the outcome measure. Associations with socioeconomic, health- and health behavior–related, psychological, and COVID-19–related determinants were tested using random effects logistic regressions.

Results: In our sample, 49% (2673/5456) of participants were female and the mean age of the participants was 67.8 (SD 9.4) years. Past experience with online health consultations was reported by 10.3% (561/5456) of the sample. Online health consultation use was associated with high education (OR 1.43, 95% CI 1.06 - 1.93; *P*=.02), poor self-rated health (OR 0.60, 95% CI 0.49 - 0.75; *P*<.001), and higher frequency of physical activity (reference: low frequency; medium frequency: OR 1.58, 95% CI 1.15 - 2.17; *P*=.005; high frequency: OR 1.73, 95% CI 1.09 - 2.76; *P*=.02). Moreover, greater levels of loneliness (OR 1.43, 95% CI 1.06 - 1.93; *P*=.04) and life satisfaction (OR 1.33, 95% CI 1.02 - 1.73; *P*=.04) as well as perceiving the COVID-19 crisis as a greater personal threat (OR 1.08, 95% CI 1.01 - 1.15; *P*=.02) were associated with having online health consultations during the COVID-19 pandemic.

Conclusions: Online health consultation use does not seem to be exclusively associated with the health of middle-aged and older patients. Study findings emphasize the longitudinal association of education and psychosocial factors as well as health factors with telemedicine use during the COVID-19 pandemic in Germany. This knowledge may help to improve and adapt services to this patient group, which could contribute to higher utilization rates in the future. Future studies are needed to verify these initial findings under postpandemic circumstances and across different countries.

(JMIR Aging 2025;8:e60311) doi:10.2196/60311

KEYWORDS

telemedicine; digital health; remote consultation; health services for the aged; patient acceptance of health care; COVID-19

Introduction

Worldwide, health care systems are facing multiple challenges in the future related to the delivery of services (eg, access and continuity of care), human resources (eg, staff distribution and sufficiency), as well as leadership and governance (eg, strategic policies) [1]. In particular, the increasing prevalence of noncommunicable diseases, disability, and multimorbidity in connection with population aging calls for new solutions in health care (eg, [2]). According to predictions of the World Health Organization (WHO) [3], the number of global deaths due to noncommunicable diseases may increase from 36 million (2008) up to 52 million in 2030.

Promising methods for the delivery of health care in the future include digital solutions such as telemedicine [4]. The WHO Global Observatory for eHealth [5] defines telemedicine as

follows: (1) its purpose is to provide clinical support, (2) it is intended to overcome geographical barriers, connecting users who are not in the same physical location, (3) it involves the use of various types of information and communication technology, and (4) its goal is to improve health outcomes. Major strengths of telemedicine include improved access to care and information, time and cost savings, as well as convenience and flexibility [6,7]. Multiple studies evaluated the effectiveness and cost-effectiveness of telemedicine services. The services seem to be effective and can produce at least comparable effects to in-person services [8,9]. Moreover, the current literature suggests that telemedicine can be a cost-effective service delivery method [8-11]. In addition, providers as well as patients seem to be highly satisfied with the services [8,9,12].

Despite its benefits, telemedicine implementations were only limited and remained at a low level up until the occurrence of the COVID-19 pandemic [13-15]. To prevent further spreading of the virus and relieve the health care system, major changes in the delivery of health care services had to be made in response to the pandemic [15]. Telemedicine represented a valuable tool to avoid personal contact during health consultations. Consequently, the global telemedicine utilization increased tremendously [13-16]. For example, Koonin et al [17] reported a 154% increase in telehealth visits in the United States in March 2020 compared with March 2019. Also in Germany, which is the focus of our study, the proportion of contract physicians and psychotherapists who offered and billed for telemedicine services increased from 6.1% (2019) to 24.6% (2021) [18]. Particularly among the psychotherapeutic care sector, utilization rates were high in Germany [18]. Since many nonessential in-person medical appointments were canceled or postponed due to the pandemic in Germany [19,20], remote health consultations represented a valuable alternative format to assure the continuity of care in spite of pandemic circumstances. This digital transformation of the German health care system was facilitated by the introduction of laws including the Digital Health Care Act [21] and the Digital Health Application Ordinance [22], which supported the prescription of health care apps, provision of video consultations, and integration of digital provider networks. In the German health care system, health insurance is compulsory with about 90% of the population having statutory and 10% having private health insurance [23]. The implementation of the new laws enabled the coverage of digital health care by statutory insurances, ensuring that telemedicine users incur no additional costs.

Considering population aging, telemedicine is particularly relevant for the older patient groups. Telemedicine was found to be effective in older adults [24-26]. For instance, van den Berg et al [25] reviewed 68 interventional telemedical studies with a controlled design examining older adults and found that none of the included studies reported better outcomes for the control group (eg, randomized or matched control groups that received usual in-person care). Moreover, older adults seem to accept and are satisfied with the services [24,27]. Taking older adults' higher need for health care and probable mobility restrictions into account, older adults can particularly benefit from the remote format. Nevertheless, older patients use telemedicine services less often than younger age groups and report multiple barriers to using the services (eg, [28-32]). For example, Wilson et al [30] identified factors such as physical difficulties (eg, visual or hearing impairments), privacy concerns, lack of experience, training, or support as barriers to the use of eHealth by older adults.

In Germany, older patients also used telemedicine services less frequently than younger patients did during the pandemic [18,33-36]. However, increasing telemedicine use in older age groups can be helpful for the future delivery of health care services in Germany to deal with problems such as physician shortages or increased demand for (long-term) care caused by demographic change [37]. Therefore, it is important to examine factors that are associated with telemedicine use in older adults to increase utilization in Germany. Previous international reviews highlighted older patients' characteristics associated with higher telemedicine use [38-40]. This included younger age, higher education, higher self-efficacy, greater experience or skills in using electronic devices, access to technology and internet, greater social support or influence, higher (health-related) motivation, and greater openness to experience as well as fewer privacy concerns and less severe health impairments.

So far, only limited evidence from Germany regarding the determinants of telemedicine use during the pandemic exists. The previous studies stressed the association of socioeconomic (male or female sex, younger age, high or low education, social status, living in an urban area, and having children under 18 years), psychosocial (loneliness, digital literacy), health (mental or physical health problems), and COVID-19–related factors (higher perceived severity of COVID-19 infection, having had COVID-19 infection, subjective COVID-19–related challenges, COVID-19–related cognitive preoccupation, anxiety, and worries) with telemedicine use [18,34-36,41-43]. Very few studies exclusively looked at middle-aged and older adults in Germany during the pandemic. These studies highlighted the positive association of telemedicine use with education, living with a partner in the same household, mental (ie, anxiety and depression) or physical health problems, loneliness, life satisfaction as well as forgoing medical treatment due to the fear of being infected by the coronavirus [44,45]. A large variety of determinants was observed in the existing studies and more research is needed to further explore and verify the findings.

In addition to the small number of studies that observe determinants of telemedicine use in middle-aged and older adults during the pandemic in Germany, hardly any studies examined the determinants longitudinally. Solely, von der Groeben et al [35] used a quasi-longitudinal design (ie, they observed cross-sectional data from 3 different time points) to detect determinants of patient use and attitude toward using video and telephone conferences in a population-representative sample of adults (18 - 69 y) affected by depression during 3 different pandemic time points in Germany. Since telemedicine use and acceptance varied over the course of the pandemic (eg, [31,35,46]), it may be beneficial to consider more than just one time point when observing telemedicine use behavior during the pandemic. Moreover, the longitudinal approach gives further insight into the directionality of the relationships. Therefore,

our study aimed to longitudinally investigate determinants of online health consultation use in a large representative sample of middle-aged and older individuals during the COVID-19 pandemic in Germany.

Expanding the knowledge regarding determinants of online health consultation use in middle-aged and older adults could help to identify target groups for telemedicine services, as well as groups that would benefit from additional support for using the services. Furthermore, this knowledge may help to adapt telemedicine services to the needs and preferences of middle-aged and older adults. Consequently, important practical and theoretical implications may be derived from our findings, which could foster greater use of telemedicine services among middle-aged and older individuals, ultimately helping to deal with future health care challenges posed by population aging.

Methods

Sample

Nationally representative cross-sectional and longitudinal data were taken from the German Ageing Survey (DEAS [47]). The DEAS focuses on the German middle-aged and older population (starting at 40 y) and aims to describe living conditions and diversity among this population as well as aging and social change processes that are related to this life stage [47]. The first wave of the DEAS was conducted in 1996, followed by further waves in 2002, 2008, 2011, 2014, 2017, 2020/2021. The survey has a cohort-sequential design in which new baseline samples were added in 2002, 2008, and 2014. The baseline samples of the DEAS were disproportionally stratified into age groups, gender, and region [47]. In response to the COVID-19 pandemic, an additional Compact Survey to measure the pandemic impact on middle- and older adults' lives in Germany was implemented in 2020.

For the purpose of our study, data from the Compact Survey [48] and wave 7 [49] of the DEAS were observed. The Compact Survey was conducted from June until July 2020 and consisted of paper-and-pencil questionnaires that were sent to individuals who had taken part in the DEAS at least once in the past. The response rate for the Compact Survey was 57% (4823/8533) [50]. Due to the ongoing pandemic during wave 7, the usual computer-assisted personal interviews were replaced by telephone-administered interviews as well as an additional paper-and-pencil questionnaire. The data collection for wave 7 took place from November 2020 until March 2021 and the response rate was 66% (5402/8207) [51]. Overall, 4103 individuals participated in both surveys [51]. Since the last DEAS baseline sample was added in 2014, participants of the Compact Survey and wave 7 of the DEAS were at least 46 years old.

The DEAS is funded by the German Federal Ministry for Family Affairs, Senior Citizens, Women, and Youth and was conducted and developed by the German Centre of Gerontology. The fieldwork was carried out by the INFAS Institute for Applied Social Sciences.

Ethical Considerations

Written informed consent was obtained from all individuals who participated in the DEAS, and opting out of the survey was possible at all times [47]. DEAS participants received an incentive (eg, €10 in DEAS wave 7; conversion rate US $1=€0.951 in 2022 [51]). Due to data protection guidelines, a data distribution contract needs to be signed prior to using the anonymized DEAS data [47]. The DEAS study complies with the Declaration of Helsinki and did not require further ethical examination since the criteria for the need of ethical approval were not met for this survey (eg, missing information regarding the study or aim of the study, examination of vulnerable groups or patients, high risk or burden for participants due to participation).

Dependent Variable

For the sake of our study, we exclusively included DEAS participants, who indicated having access to the internet (Compact survey: 3858/4676, 83%; wave 7: 3676/4276, 86%). In both DEAS waves, participants were asked "How often do you use the Internet for the following purposes?." Among the listed items was "Consultations with doctors and therapists via an online platform" in the Compact survey and "Providing consultations with doctors or therapists on online platforms" in wave 7 of the DEAS. The response format consisted of a 6-point Likert scale indicating use frequency as "never," "less often or seldom," "1 to 3 times a month," "once a week," "several times a week," or "daily." Since only a small number of participants had multiple consultations with doctors or therapists on online platforms, the outcome was dichotomized for our analysis (0="never"; 1="less often or seldom," "1 to 3 times a month," "once a week," "several times a week," or "daily").

Independent Variables

Based on theoretical considerations and previous research [38-40,44,45], different groups of determinants were considered. This included socioeconomic, health- and health behavior–related, psychological, and COVID-19–related determinants. Regarding socioeconomic characteristics, we examined sex, age, educational level (International Standard Classification of Education 97 [52]: low, medium, or high education), employment status (used, retired, and other or unemployed), household income, migration background (no, yes), area lived in (metropolitan districts, urban districts, [partially] densely populated rural districts, sparsely populated rural districts), residential form of partnership (no partner, partner in the same household, and partner not in the same household), and presence of children (no, yes).

Health-related factors included self-rated health (ranging from 0=very bad to 4=very good) as well as the frequency of physical activity and walks (ranging from 0=never to 5=daily) as measures of health-related behavior. For the sake of our analysis, the values for the frequency of physical activity and walks were divided into tertiles and were included as categorical variables (low, medium, and high frequency).

Psychological determinants that were considered were depressive symptoms, loneliness, attitude toward own aging, and life satisfaction. Depressive symptoms were measured using

a 10-item German short form of the Center for Epidemiologic Studies - Depression scale [53] (CES-D; scores ranging from 0 to 30, higher values indicate more severe depressive symptoms) in the Compact survey. In wave 7 of the DEAS, depressive symptoms were measured using the German version of the 15-item CES-D [54] (scores ranging from 0 to 45, higher values indicate more severe depressive symptoms). The values for both scales were standardized to assure comparability between both surveys in our analysis. Both of these well-established instruments were evaluated in the past and have good psychometric properties [53,55,56]. Cronbach α values for both scales were 0.83 (Compact survey) and 0.84 (wave 7) and McDonald Omega was 0.85 (Compact survey) and 0.86 (wave 7) in our sample. In addition, loneliness was measured with the 6-item De Jong Gierveld Loneliness Scale [57] (scores ranging from 1 to 4, higher values indicate higher levels of loneliness). The scale has favorable psychometric properties [57] (Compact survey: Cronbach α=0.78, McDonald's Omega =0.79; wave 7: Cronbach α=0.80, McDonald Omega =0.81). The self-perception of one's own aging was examined using the German version of the 5-item Attitude Toward Own Aging subscale of the Philadelphia Geriatric Center Morale Scale [58,59] (scores ranging from 1 to 4, higher values indicate a more positive perception of own aging). This widely used scale was evaluated in different age groups in the past (eg, [60]). In our sample, Cronbach α was 0.77 (Compact Survey and wave 7) and McDonald Omega was 0.77 (Compact Survey and wave 7). Life satisfaction was measured with the German version of the Satisfaction with Life Scale [61,62] (scores ranging from 1 to 5, higher values indicate greater life satisfaction). The German version of the scale was evaluated in the past and showed good psychometric properties [63]. For this scale, the Cronbach α was 0.86 (Compact survey) and 0.84 (wave 7) and the McDonald Omega was 0.87 (Compact survey) and 0.86 (wave 7) in our sample.

Finally, we controlled for COVID-19–related determinants, which included perceiving the Corona crisis as a personal threat (scores ranging from 1=not at all a threat for me to 10=extreme threat for me), past infection with the Coronavirus by oneself (no, yes, and unknown), or by people from one's personal environment (no, yes, and unknown) as well as the feeling of being able to influence the infection with the Coronavirus (scores ranging from 1=not at all to 7=entirely).

Statistical Analysis

In the first step, sample characteristics of our pooled analytic sample were computed. The analytic sample consisted of individuals who participated in at least one of the 2 surveys (5456 observations corresponding to 3222 individuals). Second, random effects logistic regressions were calculated to test the associations of the determinants with online health consultation use. The random effects regression model considers the panel structure of the data and allows the inclusion of not only time-varying but also time-constant predictors in our model, under the assumption that unobserved unit-specific heterogeneity is not correlated with the independent variables [64]. When this independence assumption is fulfilled, the random effects model may be more efficient than the fixed effects model as it considers both between and within variation [64]. Our choice was supported by the Hausman test. The null hypothesis of the Hausman test states that both models (fixed and random effects model) are consistent while the random effects model is more efficient [64,65]. Therefore, the random effects model is preferred when the null hypothesis cannot be rejected. Since the Hausman test statistic was nonsignificant for our sample (P=.72), we used random effects models for our analysis. Stata (version 16.0, StataCorp) was used for the statistical analyses and the random effects logistic regression was calculated using the "xtlogit" command with the "re" option. The sample was stratified by sex and age groups (≤64 and ≥65 years) in additional analyses. Statistical significance was defined as an alpha level of P<.05. Missing data were handled using listwise deletion.

Results

Sample Characteristics

The pooled analytic sample characteristics for all included variables are presented in Table 1. In the pooled sample of the Compact survey and wave 7 of the DEAS, 49% (2673/5456) were female and the mean age of the participants was 67.8 (SD 9.4) years. When examining past consultations with doctors or therapists on online platforms, 10.3% (561/5456) reported past experience with online consultations.

Table . Analytic pooled sample characteristics (N=5456).

Characteristics[a]	Values
Consultations with doctors or therapists on online platforms, n (%)	
No	4895 (89.7)
Yes	561 (10.3)
Sex, n (%)	
Male	2783 (51)
Female	2673 (49)
Age (years), mean (SD)	67.8 (9.4)
Educational level, n (%)	
Low (ISCED[b] 0 - 2) or medium (ISCED 3 - 4)	2455 (45)
High (ISCED 5 - 6)	3001 (55.0)
Employment status, n (%)	
Employed	1745 (32.0)
Retired	3475 (63.7)
Other or unemployed	236 (4.3)
Monthly household income (€), mean (SD)[c]	4051.2 (11,806.1)
Migration background, n (%)	
No	5223 (95.7)
Yes	233 (4.3)
Area lived in, n (%)	
Metropolitan districts	1524 (27.9)
Urban districts	2014 (36.9)
(Partially) densely populated rural districts	1135 (20.8)
Sparsely populated rural districts	783 (14.4)
Residential form of partnership, n (%)	
No partner	1069 (19.6)
Partner in the same household	4152 (76.1)
Partner not in the same household	235 (4.3)
Having children, n (%)	
None	602 (11)
One or more	4854 (89)
Self-rated health, mean (SD)[d]	2.6 (0.8)
Frequency of physical activity, n (%)	
Low frequency	1589 (29.1)
Medium frequency	3234 (59.3)
High frequency	633 (11.6)
Frequency of walks, n (%)	
Low frequency	2135 (39.1)
Medium frequency	2147 (39.4)
High frequency	1174 (21.5)
Depressive symptoms, mean (SD)[e]	−0.1 (0.9)
Loneliness, mean (SD)[f]	1.8 (0.5)

Characteristics[a]	Values
Life satisfaction, mean (SD)[g]	3.9 (0.7)
Attitude toward own aging, mean (SD)[h]	3.0 (0.5)
Perceiving the Corona crisis as a personal threat, mean (SD)[i]	4.3 (2.1)
Oneself infected with the Coronavirus, n (%)	
No	5244 (96.1)
Yes	57 (1)
Unknown	155 (2.8)
People from personal environment infected with the Coronavirus, n (%)	
No	4517 (82.8)
Yes	853 (15.6)
Unknown	86 (1.6)
Feeling that one can influence the infection with the Coronavirus, mean (SD)[j]	4.7 (1.4)

[a]Due to differences in measurement tools, values for depressive symptoms had to be standardized. The standardized values should be interpreted as number of SDs by which the original values lay above or below their mean. For example, a value of 1 or -1 indicates that the reported overall score in the CES-D lays one SD above/below the mean CES-D score.

[b]ISCED: International Standard Classification of Education.

[c] Range (0-500,000); the conversion rate of USD to Euro was US $1=€0.846 in 2021 and US $1=€0.951 in 2022.

[d]Range 0-4. Higher values indicate better self-rated health.

[e]Range for standardized values –1.8 to 5.9. Higher values indicate more depressive symptoms.

[f]Range 1-4. Higher values indicate higher loneliness levels.

[g]Range 1-5. Higher values indicate greater life satisfaction.

[h]Range 1-4. Higher values indicate a more positive perception of own aging.

[i]Range 1=not at all a threat for me to 10=extreme threat for me.

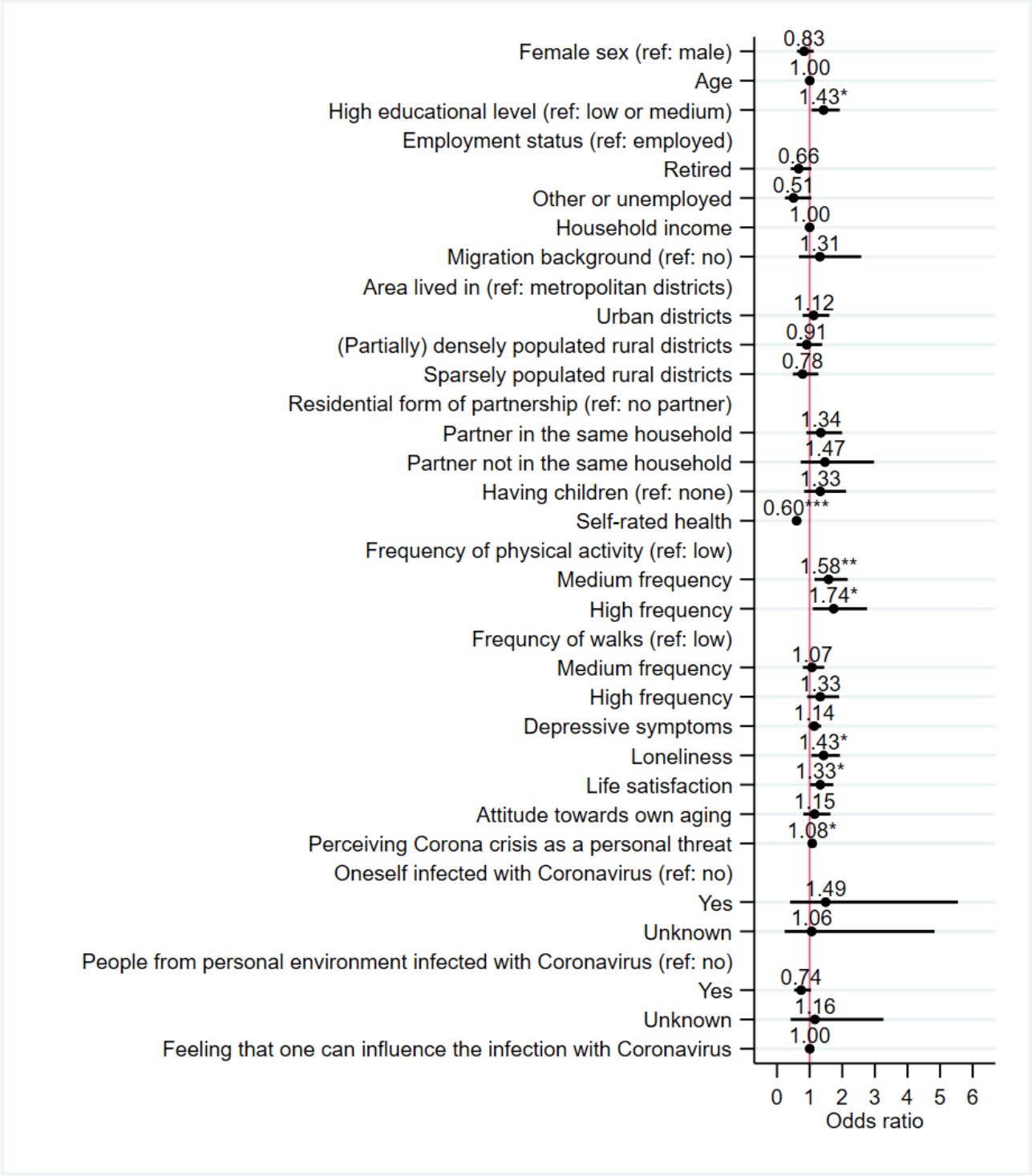
[j]Range 1=not at all to 7=entirely.

Regression Analysis

The results of the random effects logistic regression are presented in Figure 1 (Table S1 in Multimedia Appendix 1). The majority of the determinants were not significantly associated with online health consultation use during the COVID-19 pandemic. Nevertheless, we found a significant longitudinal association of the outcome with high education (OR 1.43, 95% CI 1.06 - 1.93; *P*=.02), poor self-rated health (OR 0.60, 95% CI 0.49 - 0.75; *P*<.001), higher frequency of physical activity (medium frequency: OR 1.58, 95% CI 1.15 - 2.17; *P*=.005; high frequency: OR 1.73, 95% CI 1.09 - 2.76; *P*=.02), higher loneliness (OR 1.43, 95% CI 1.06 - 1.93; *P*=.04), greater life satisfaction (OR 1.33, 95% CI 1.02 - 1.73; *P*=.04), and perceiving the Corona crisis as a greater personal threat (OR 1.08, 95% CI 1.01 - 1.15; *P*=.02).

In additional analyses, we stratified the sample by sex and age group (Tables S2 and S3 in Multimedia Appendix 1 for more details). While online health consultation use in female participants was only associated with poor self-rated health (OR 0.62, 95% CI 0.45 - 0.85; *P*=.004), in male individuals the outcome was associated with poor self-rated health (OR 0.59, 95% CI 0.44 - 0.79; *P*<.001) as well as high education (OR 1.70, 95% CI 1.12 - 2.56; *P*=.01), living with a partner in the same household (OR 2.27, 95% CI 1.13 - 4.59; *P*=.02), living with a partner not in the same household (OR 2.86, 95% CI 1.04 - 7.87; *P*=.04), higher loneliness (OR 1.94, 95% CI 1.25 - 2.99; *P*=.003), higher frequency of physical activity (medium frequency: OR 1.88, 95% CI 1.22 - 2.90; *P*=.004; high frequency: OR 2.17, 95% CI 1.17 - 4.03; *P*=.01) and walks (high frequency: OR 1.99, 95% CI 1.24 - 3.20; *P*=.004). In participants aged ≤64 years, online health consultation use was associated with poor self-rated health (OR 0.59, 95% CI 0.42 - 0.83; *P*=.003) and higher frequency of physical activity (medium frequency: OR 1.91, 95% CI 1.14 - 3.21; *P*=.01). In older participants (≥65 years), poor self-rated health (OR 0.61, 95% CI 0.46 - 0.81; *P*=.001) as well as older age (OR 1.03, 95% CI 1.00 - 1.07; *P*=.05), high education (OR 1.58, 95% CI 1.06 - 2.37; *P*=.026), and higher loneliness (OR 1.52, 95% CI 1.01 - 2.29; *P*=.04) were associated with online health consultation use.

Figure 1. Results of random effects logistic regression for determinants of online health consultation use during the COVID-19 pandemic (N=5456). Odds ratios with 95% CI are reported. Unless stated otherwise, the reference category is always zero or absence of the characteristic. ref: reference category; *** *P*<.001; ** *P*<.01; * *P*<.05.



Discussion

Principal Findings

Nationally representative longitudinal data from Germany were used to observe the longitudinal association of various determinants with the use of online consultations with doctors or therapists during the COVID-19 pandemic in a large sample of middle-aged and older individuals with access to the internet. Random effects logistic regressions revealed associations of education, health, and psychosocial factors with online health consultation use during the pandemic. In additional analyses stratified by sex and age, self-rated health was negatively associated with the outcome in all groups. Additional relationships with age, education, relationship status, and loneliness were only observed among the male and older (≥65 years) subgroups. Considering the limited evidence regarding determinants of telemedicine use (particularly based on longitudinal data), our longitudinal study considerably extend current knowledge on socioeconomic, health- and health

behavior–related, psychological, and COVID-19-related determinants.

Relation to Previous Research

In contrast to findings from other German samples [18,33,35,36,41,42,66], most of the socioeconomic determinants, such as sex or age, were not associated with online health consultation use in our sample. Nevertheless, a study that exclusively observed an older German sample also found no associations with socioeconomic characteristics [45]. This could imply that socioeconomic characteristics are less relevant for telemedicine use in middle-aged and older patients in Germany. However, the individual's educational level was positively associated with the outcome in our analysis. This is in line with findings from other large German [33,34,36,42] and international samples [38,40,67,68] and could suggest that high education is linked to higher digital literacy as well as access to necessary technical equipment for using telemedicine. A systematic review by Estrela et al [69] examined 36 international articles on digital health literacy and highlighted its positive relationship with education. Correspondingly, in a large US sample of old-aged Medicare beneficiaries (≥65 years), Choi et al [70] found associations of younger age and higher income with telemedicine use during the COVID-19 pandemic, which disappeared when controling for technology-enabling factors (eg, information and communication technology device ownership or use experience). Consequently, technology-enabling factors could be particularly considered when trying to enhance telemedicine use among older patients. For instance, implementing supporting material (eg, leaflets or video instructions), technical support (eg, via telephone) or hybrid care combinations in clinical practice may be helpful to strengthen competence and digital health literacy in older patients.

Health and health behavior were associated with online health consultation use in our sample. Individuals who reported a poorer health status were more likely to use online health consultations during the pandemic. Likewise, severe health limitations or poor health were associated with telemedicine use in the German Survey of Health, Ageing and Retirement in Europe (SHARE) sample of middle-aged and older adults [45] and in international samples [68,71]. This association could have been caused by greater health needs in unhealthy individuals or may be connected to precautions due to the COVID-19 pandemic. Likewise, perceiving the Corona crisis as a greater personal threat was associated with online health consultation use. Telemedicine might have presented a treatment option for patients who were scared of becoming infected with the Coronavirus and wanted to avoid personal contact (eg, [72]). Fear of the Coronavirus and pandemic-related challenges were also associated with telemedicine use in other German samples during the pandemic [36,41,42,45]. Other COVID-19–related factors such as the infection of oneself or close others with the virus were not associated with the outcome in our sample. A reason for that could be that only a small proportion of our pooled sample was infected with the virus (1%) or knew someone in their close personal environment who was infected (15.6%). Furthermore, the frequency of physical activity was positively associated with online consultation use. Physical activity is an important determinant of health and was associated with higher utilization of preventive or office-based health services and lower use of inpatient or emergency care among adults in previous international research (eg, [73,74]). Therefore, physically active individuals seem to show greater levels of health awareness, which might have been connected to the higher telemedicine use in this group.

Some psychological factors were associated with the outcome in our sample. Whereas depressive symptoms and the attitude towards one's own aging did not show an association, psychosocial factors including loneliness and life satisfaction had a significant relationship with online health consultation use. In the German SHARE sample, a positive association of depressive symptoms as well as loneliness with telemedicine use in middle-aged and older individuals was observed [45]. The mixed evidence concerning the relationship of depressive symptoms with telemedicine utilization might be explained by the different pandemic periods that were considered in the studies (Summer and Winter 2020 vs Summer 2021). In fact, depressive symptoms [75] and telemedicine acceptance [35] were found to have increased over the course of the pandemic. Regarding loneliness, Robbins et al [76] also observed higher loneliness rates among telemedicine users (telephone contacts), while in-person visits were associated with fewer feelings of loneliness among older adults (≥65 years) residing in the United States during the pandemic. It might be the case that older adults who indicated higher levels of loneliness were more open to using telemedicine services to satisfy their unmet social needs during pandemic times. Telemedicine services may be more accessible for older individuals (eg, no traveling for mobility-restricted individuals needed), which might encourage lonely individuals to take the initiative to foster social interaction through telemedicine appointments. Regarding life satisfaction, König et al [77] recently observed a positive correlation between life satisfaction and digital health literacy in a nationally representative survey of the population in Germany, which could explain the observed relationship between higher life satisfaction and online health consultation use in our sample. In addition, higher life satisfaction was associated with health-promoting behaviors in previous studies [78-80], which could have contributed to higher online health consultation use.

Additional analyses stratified by sex and age further highlighted the relationships with the determinants, especially in male and older participants. While online consultation use in females was only associated with health needs, male individual's use behavior was additionally associated with psychosocial factors (ie, relationship status and loneliness), education, and health behavior (ie, frequency of walks and physical activity). The decision of male patients to use telemedicine services seems to be connected to additional factors and therefore more complex compared with female patients. Male patients may be facing additional barriers to telemedicine utilization. Therefore, future research is needed to explore gender-specific determinants and barriers to the use of telemedicine. Furthermore, particularly among individuals aged 65 years and older, loneliness, education, and age were associated with online health consultation use. Consequently, compared with individuals aged

40 to 64 years, older patients might be especially affected by disparities in education or loneliness.

When considering the international context, multiple reviews mainly based on quantitative cross-sectional, randomized controlled or qualitative studies observed determinants of telemedicine use in older age groups and found positive associations with educational level, health needs, and health-related motivation to use the services [38-40], which is in line with our findings. Moreover, they observed additional relationships (eg, with age, sex, and social support or influence), which we did not find in our analysis. Age or sex were not associated with overall telemedicine use in our sample. The reviews [38-40] mostly included studies from the United States or Europe. Telemedicine regulations differ substantially between Germany and the United States, but also in the European context. Whereas Germany has implemented national telemedicine coverage rules during the pandemic, large state-specific variation in insurance coverage and regulations exist in the United States [71,81], which may have caused additional barriers for older telemedicine users. Nevertheless, our stratified analyses indicated differences in determinants of use in the different sex and age groups. Furthermore, we observed a positive relationship of loneliness with telemedicine use, which is in contrast to the observed negative relationship of social isolation or lack of social support with telemedicine use in international studies [39]. Raja et al [82] reviewed studies of older adults in European countries and found that social support and lack of social support were both associated with using new technologies, including telemedicine. Social support may be crucial when older adults are faced with problems when learning new technologies, which was also observed in US samples [68]. Nevertheless, lack of social support or loneliness might also motivate older adults to try out new technologies to address their social needs, which we observed in our sample. Moreover, the discrepancies might also be explained by differences in regulations or access to telemedicine care (eg, variations in out-of-pocket payments, supply, or complexity of use). For instance, the requirement for additional out-of-pocket payments or high barriers to use may lower the probability of using telemedicine services for the primary purpose of social interaction. Future research concerning differences in psychosocial determinants across different countries is needed.

Strengths and Limitations

The nationally representative, large DEAS sample of middle-aged and older individuals in Germany represents a key strength of our study. Middle-aged and older adults are a major target group for future telemedicine services in Germany, thus it is of particular importance to explore the telemedicine use behavior of this age group. In addition, longitudinal data were exploited, which enabled us to consider two different pandemic stages and account for the exceptional circumstances during that time. Since only few studies observed determinants of online health consultation use in German middle-aged and older adults in the past, our study adds valuable knowledge to the existing literature.

Nevertheless, some limitations should be noted. Telemedicine use was represented by having online consultations with doctors or therapists in our study. We neither examined specific patient groups nor focused on a certain telemedicine format (eg, video conferences or mobile apps). Therefore, we observed a potentially heterogeneous user group. Since telemedicine acceptance during the pandemic was found to vary across medical specialties and telemedicine formats [31], future research that tests for differences in use among patient groups or different telemedicine formats is needed to tailor future services to major user groups. However, our study provides initial insights into telemedicine use in middle and old age. In addition, the DEAS panel holds a slight selection bias. Young, highly educated, healthier, and female individuals were somewhat more likely to participate in the DEAS [83]. However, selection bias in the DEAS sample was found to be small and the distribution of major sociodemographic characteristics closely mirrors the distribution within the overall population of Germany [47]. Moreover, only individuals with access to the internet were included in our study. Therefore, generalization of the results might be slightly limited for some groups of the German middle-aged and older population.

Conclusion

Telemedicine services represent a valuable tool to deal with the increasing demand for health care caused by population aging. Knowledge about telemedicine use and its determinants, particularly in middle-aged and older individuals, is essential to promote widespread implementations in the future. Our study highlights the relationship of education, psychosocial, and health factors with telemedicine use of community-dwelling middle-aged and older individuals in Germany during the COVID-19 pandemic. Therefore, telemedicine use does not only depend on health needs of middle-aged and older patients. The finding that particularly highly educated individuals used online health consultations may point toward social inequality among telemedicine users. Consequently, efforts should be made to enable access to telemedicine for all patient groups and individual support should be provided (eg, for patients with low [digital] health literacy) to remove barriers to telemedicine use. Moreover, special attention should be paid to individuals with low life satisfaction and an unhealthy lifestyle since they seem harder to reach through telemedicine services. Finally, future research is needed to test the relevance of the observed relationships in the postpandemic context and identify potential reasons for use or nonuse of telemedicine services in middle-aged and older adults in Germany (eg, based on qualitative data). Moreover, cross-country comparisons regarding the determinants of telemedicine use remain to be explored.

Acknowledgments

We acknowledge financial support from the Open Access Publication Fund of UKE - Universitätsklinikum Hamburg-Eppendorf.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of random effects logistic regression for determinants of online health consultation use during the COVID-19 pandemic and additional analyses stratified by sex and age.
[DOCX File, 34 KB - aging_v8i1e60311_app1.docx]

References

1. Lehoux P, Roncarolo F, Silva HP, Boivin A, Denis JL, Hébert R. What health system challenges should responsible innovation in health address? Insights from an international scoping review. Int J Health Policy Manag 2019 Feb 1;8(2):63-75. [doi: 10.15171/ijhpm.2018.110] [Medline: 30980619]
2. Mahishale V. Ageing world: health care challenges. J Sci Soc 2015;42(3):138. [doi: 10.4103/0974-5009.165540]
3. Global status report on noncommunicable diseases 2010. World Health Organization. 2011. URL: https://iris.who.int/bitstream/handle/10665/44579/9789240686458_eng.pdf?sequence=1&isAllowed=y [accessed 2025-05-20]
4. Khemapech T. Telemedicine – meaning, challenges and opportunities. Siriraj Med Bull 2019 May 1;71(3):246-252. [doi: 10.33192/Smj.2019.38]
5. Telemedicine: opportunities and developments in member states: report on the second global survey on ehealth. WHO Global Observatory for eHealth. 2010. URL: https://iris.who.int/bitstream/handle/10665/44497/9789241564144_eng.pdf?sequence=1&isAllowed=y [accessed 2025-05-22]
6. Hjelm NM. Benefits and drawbacks of telemedicine. J Telemed Telecare 2005;11(2):60-70. [doi: 10.1258/1357633053499886] [Medline: 15829049]
7. Meneses ADF, Pimentel FF, da Cruz JPF, Candido Dos Reis FJ. Experiences of women with breast cancer using telehealth: a qualitative systematic review. Clin Breast Cancer 2023 Feb;23(2):101-107. [doi: 10.1016/j.clbc.2022.11.001] [Medline: 36464603]
8. Goharinejad S, Hajesmaeel-Gohari S, Jannati N, Goharinejad S, Bahaadinbeigy K. Review of systematic reviews in the field of telemedicine. Med J Islam Repub Iran 2021;35(1):184. [doi: 10.47176/mjiri.35.184] [Medline: 36042824]
9. Eze ND, Mateus C, Cravo Oliveira Hashiguchi T. Telemedicine in the OECD: An umbrella review of clinical and cost-effectiveness, patient experience and implementation. PLoS ONE 2020;15(8):e0237585. [doi: 10.1371/journal.pone.0237585] [Medline: 32790752]
10. Kichloo A, Albosta M, Dettloff K, et al. Telemedicine, the current COVID-19 pandemic and the future: a narrative review and perspectives moving forward in the USA. Fam Med Community Health 2020 Aug;8(3):e000530. [doi: 10.1136/fmch-2020-000530] [Medline: 32816942]
11. Atmojo JT, Sudaryanto WT, Widiyanto A, et al. Telemedicine, cost effectiveness, and patients satisfaction: a systematic review. J Health Policy Manage 2020;5(2):103-107 [FREE Full text] [doi: 10.26911/thejhpm.2020.05.02.02]
12. Nguyen M, Waller M, Pandya A, Portnoy J. A review of patient and provider satisfaction with telemedicine. Curr Allergy Asthma Rep 2020 Sep 22;20(11):72. [doi: 10.1007/s11882-020-00969-7] [Medline: 32959158]
13. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
14. Friedman AB, Gervasi S, Song H, et al. Telemedicine catches on: changes in the utilization of telemedicine services during the COVID-19 pandemic. Am J Manag Care 2022 Jan 1;28(1):e1-e6. [doi: 10.37765/ajmc.2022.88771] [Medline: 35049260]
15. Seivert S, Badowski ME. The rise of telemedicine: lessons from a global pandemic. EMJ Innov 2021;5(1):64-69. [doi: 10.33590/emjinnov/20-00239]
16. Nittari G, Savva D, Tomassoni D, Tayebati SK, Amenta F. Telemedicine in the COVID-19 era: a narrative review based on current evidence. Int J Environ Res Public Health 2022 Apr 22;19(9):5101. [doi: 10.3390/ijerph19095101] [Medline: 35564494]
17. Koonin LM, Hoots B, Tsang CA, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599. [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
18. Heuer J, Osterwald A, Akmatov MK, Zentralinstitut für die kassenärztliche Versorgung in Deutschland (Zi), et al. Telemedizin als alternativer Zugang zu vertragsärztlicher ambulanter Versorgung–Trends im Zeitraum 2017 bis 2021 [Article in German]. Versorgungsatlas-Bericht Nr 2023;23(6):1-21. [doi: 10.20364/VA-23.06]
19. Heidemann C, Reitzle L, Schmidt C, Fuchs J, Prütz F, Scheidt-Nave C. Non-utilisation of health care services during the COVID-19 pandemic: results of the CoMoLo study. J Health Monit 2022 Mar;7(Suppl 1):2-17. [doi: 10.25646/9564] [Medline: 35356069]
20. Schäfer I, Haack A, Neumann M, Koch-Gromus U, Scherer M, Petersen E. Non-utilisation of medical services during the COVID-19 pandemic among persons with chronic diseases. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2023 Mar;66(3):275-282. [doi: 10.1007/s00103-023-03665-9] [Medline: 36723636]

21. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz—DVG) [digital healthcare act]. BGBI I 2562. 2019. URL: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&start=%2F%2F%2A%5B%40attr_id=%27bgbl119s2562.pdf%27%5D#/switch/tocPane?_ts=1745927445559 [accessed 2025-05-22]
22. Verordnung über das Verfahren und die Anforderungen zur Prüfung der Erstattungsfähigkeit digitaler Gesundheitsanwendungen in der gesetzlichen Krankenversicherung (Digitale Gesundheitsanwendungen-Verordnung - DIGAV) [digital health applications ordinance]. BGBl I. 2020. URL: https://www.gesetze-im-internet.de/digav/BJNR076800020.html [accessed 2025-05-20]
23. Passon A, Lüngen M, Gerber A, Redaelli M, Stock S, In: Lauterbach K, Stock S, Brunner H, editors. Gesundheitsökonomie. Das Krankenversicherungssystem in Deutschland: Verlag Hans Huber, Hogrefe AG; 2009:105-136.
24. Şahin E, Yavuz Veizi BG, Naharci MI. Telemedicine interventions for older adults: a systematic review. J Telemed Telecare 2024 Feb;30(2):305-319. [doi: 10.1177/1357633X211058340] [Medline: 34825609]
25. van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients--a systematic review. Maturitas 2012 Oct;73(2):94-114. [doi: 10.1016/j.maturitas.2012.06.010] [Medline: 22809497]
26. Soh YY, Zhang H, Toh JJY, Li X, Wu XV. The effectiveness of tele-transitions of care interventions in high-risk older adults: a systematic review and meta-analysis. Int J Nurs Stud 2023 Mar;139:104428. [doi: 10.1016/j.ijnurstu.2022.104428] [Medline: 36682322]
27. Alsabeeha NHM, Atieh MA, Balakrishnan MS. Older adults' satisfaction with telemedicine during the COVID-19 pandemic: a systematic review. Telemed J E Health 2023 Jan;29(1):38-49. [doi: 10.1089/tmj.2022.0045] [Medline: 35613380]
28. Reinhardt G, Schwarz PEH, Harst L. Non-use of telemedicine: a scoping review. Health Informatics J 2021;27(4):14604582211043147. [doi: 10.1177/14604582211043147] [Medline: 34696613]
29. Saigí-Rubió F, Borges do Nascimento IJ, Robles N, et al. The current status of telemedicine technology use across the world health organization european region: an overview of systematic reviews. J Med Internet Res 2022 Oct 27;24(10):e40877. [doi: 10.2196/40877] [Medline: 36301602]
30. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
31. Drake C, Lian T, Cameron B, Medynskaya K, Bosworth HB, Shah K. Understanding telemedicine's "New Normal": variations in telemedicine use by specialty line and patient demographics. Telemed J E Health 2022 Jan;28(1):51-59. [doi: 10.1089/tmj.2021.0041] [Medline: 33769092]
32. Ternes S, Lavin L, Vakkalanka JP, et al. The role of increasing synchronous telehealth use during the COVID-19 pandemic on disparities in access to healthcare: A systematic review. J Telemed Telecare 2024 Apr 22;2024:1357633X241245459. [doi: 10.1177/1357633X241245459] [Medline: 38646804]
33. Reitzle L, Schmidt C, Färber F, et al. Perceived access to health care services and relevance of telemedicine during the COVID-19 pandemic in Germany. Int J Environ Res Public Health 2021 Jul 19;18(14):7661. [doi: 10.3390/ijerph18147661] [Medline: 34300110]
34. Uncovska M, Freitag B, Meister S, Fehring L. Patient acceptance of prescribed and fully reimbursed mHealth apps in Germany: an UTAUT2-based online survey study. J Med Syst 2023 Jan 27;47(1):14. [doi: 10.1007/s10916-023-01910-x] [Medline: 36705853]
35. von der Groeben S, Czaplicki A, Hegerl U, Reich H. Telemedicine during the COVID-19 pandemic in Germany: Results from three nationally representative surveys on use, attitudes and barriers among adults affected by depression. Internet Interv 2023 Apr;32:100622. [doi: 10.1016/j.invent.2023.100622] [Medline: 37091132]
36. Zöllner JP, Noda AH, McCoy J, et al. Use of health-related apps and telehealth in adults with epilepsy in Germany: a multicenter cohort study. Telemed J E Health 2023 Apr;29(4):540-550. [doi: 10.1089/tmj.2022.0238] [Medline: 35984859]
37. Kis A, Augustin M, Augustin J. Regional healthcare delivery and demographic change in Germany – scenarios for dermatological care in 2035. J Deutsche Derma Gesell 2017 Dec;15(12):1199-1209. [doi: 10.1111/ddg.13379]
38. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Sep 5;16 Suppl 5(Suppl 5):290. [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
39. Turcotte S, Bouchard C, Rousseau J, et al. Factors influencing older adults' participation in telehealth interventions for primary prevention and health promotion: A rapid review. Australas J Ageing 2024 Mar;43(1):11-30. [doi: 10.1111/ajag.13244] [Medline: 38014903]
40. Wang J, Fu Y, Lou V, Tan SY, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021 Sep;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
41. Hajek A, De Bock F, Merkel C, Kretzler B, König HH. Use of and satisfaction with telemedicine services during the pandemic: findings from the COVID-19 snapshot monitoring in Germany (COSMO). Healthcare (Basel) 2022 Jan 4;10(1):92. [doi: 10.3390/healthcare10010092] [Medline: 35052257]

42. Hannemann N, Götz NA, Schmidt L, Hübner U, Babitsch B. Patient connectivity with healthcare professionals and health insurer using digital health technologies during the COVID-19 pandemic: a German cross-sectional study. BMC Med Inform Decis Mak 2021 Aug 25;21(1):250. [doi: 10.1186/s12911-021-01605-8] [Medline: 34433452]
43. Rauschenberg C, Schick A, Goetzl C, et al. Social isolation, mental health, and use of digital interventions in youth during the COVID-19 pandemic: a nationally representative survey. Eur Psychiatry 2021 Mar 9;64(1):e20. [doi: 10.1192/j.eurpsy.2021.17] [Medline: 33686930]
44. Hajek A, König HH. Frequency and correlates of online consultations with doctors or therapists in middle-aged and older adults: nationally representative cross-sectional study. J Med Internet Res 2022 Apr 7;24(4):e29781. [doi: 10.2196/29781] [Medline: 35389360]
45. Neumann A, König HH, Hajek A. Determinants of telemedicine service use among middle-aged and older adults in Germany during the COVID-19 pandemic: cross-sectional survey study. JMIR Aging 2024 Apr 23;7:e50938. [doi: 10.2196/50938] [Medline: 38654578]
46. Mandal S, Wiesenfeld BM, Mann D, et al. Evidence for telemedicine's ongoing transformation of health care delivery since the onset of COVID-19: retrospective observational study. JMIR Form Res 2022 Oct 14;6(10):e38661. [doi: 10.2196/38661] [Medline: 36103553]
47. Klaus D, Engstler H, Mahne K, et al. Cohort profile: the German ageing survey (DEAS). Int J Epidemiol 2017 Aug 1;46(4):1105-1105g. [doi: 10.1093/ije/dyw326] [Medline: 28180273]
48. Vogel C, Wettstein M, Klaus D, Spuling S, Kortmann L, Alcántara AL, et al. German Ageing Survey: Instruments of the Compact Survey 2020: German centre of gerontology; 2021. [doi: 10.1007/978-3-319-69892-2_1115-1]
49. Vogel C, Wettstein M, Klaus D, Spuling S, Kortmann L, Alcántara AL, et al. German ageing survey, instruments of the survey 2020/2021: German centre of gerontology; 2021. [doi: 10.1007/978-3-030-22009-9_1115]
50. Engstler H. Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS2020, version 1.0 (märz 2021) [Website in German]. : Deutsches Zentrum für Altersfragen; 2021. URL: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-72029-0 [accessed 2025-05-20]
51. Stuth S. Deutscher Alterssurvey (DEAS): Kurzbeschreibung des SUF DEAS2020/21, version 11 [Website in German]. : Deutsches Zentrum für Altersfragen; 2022. URL: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-79741-0 [accessed 2025-05-20]
52. United nations educational scientific and cultural organization. International Standard Classification of Education: ISCED 1997: UNESCO Institute for Statistics 2006. URL: https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf [accessed 2025-05-23]
53. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D. Am J Prev Med 1994 Mar;10(2):77-84. [doi: 10.1016/S0749-3797(18)30622-6]
54. Hautzinger M, Bailer M. Allgemeine Depressions-Skala (ADS)[General Depression Scale]. Beltz 1993.
55. Mohebbi M, Nguyen V, McNeil JJ, et al. Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. Gen Hosp Psychiatry 2018;51:118-125. [doi: 10.1016/j.genhosppsych.2017.08.002] [Medline: 28890280]
56. Hautzinger M. Die CES-D Skala: Ein Depressionsmesinstrument für Untersuchungen in der Allgemeinbevölkerung. [The CES-D scale: a depression-rating scale for research in the general population.]. Diagnostica 1988;34(2):167-173.
57. Gierveld JDJ, Tilburg TV. A 6-item ccale for overall, emotional, and social loneliness: confirmatory tests on survey data. Res Aging 2006;28(5):582-598. [doi: 10.1177/0164027506289723]
58. Lawton MP. The Philadelphia Geriatric Center Morale Scale: a revision. J Gerontol 1975 Jan;30(1):85-89. [doi: 10.1093/geronj/30.1.85] [Medline: 1109399]
59. Liang J, Bollen KA. The structure of the Philadelphia Geriatric Center Morale scale: a reinterpretation. J Gerontol 1983 Mar;38(2):181-189. [doi: 10.1093/geronj/38.2.181] [Medline: 6827034]
60. Jung S, Siedlecki KL. Attitude toward own aging: age invariance and construct validity across middle-aged, young-old, and old-old adults. J Adult Dev 2018 Jun;25(2):141-149. [doi: 10.1007/s10804-018-9283-3]
61. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
62. Schumacher J. S. Satisfaction with life scale. In: Klaiberg A, Brähler E, editors. Diagnostische Verfahren Zu Lebensqualität Und Wohlbefinden2003.
63. Glaesmer H, Grande G, Braehler E, Roth M. The German version of the satisfaction with life scale (SWLS). Eur J Psychol Assess 2011 Jan;27(2):127-132. [doi: 10.1027/1015-5759/a000058]
64. Andreß HJ, Golsch K, Schmidt AW. Panel analysis of categorical variables. In: Andreß HJ, Golsch K, Schmidt AW, editors. Applied Panel Data Analysis for Economic and Social Surveys: Springer Berlin Heidelberg; 2013:203-285. [doi: 10.1007/978-3-642-32914-2_5]
65. Hausman JA. Specification tests in econometrics. Econometrica 1978 Nov;46(6):1251. [doi: 10.2307/1913827]
66. Muehlensiepen F, Petit P, Knitza J, Welcker M, Vuillerme N. Factors associated with telemedicine use among patients with rheumatic and musculoskeletal disease: secondary analysis of data from a German nationwide survey. J Med Internet Res 2023 Jan 27;25:e40912. [doi: 10.2196/40912] [Medline: 36705950]

67. Frydman JL, Li W, Gelfman LP, Liu B. Telemedicine uptake among older adults during the COVID-19 pandemic. Ann Intern Med 2022 Jan;175(1):145-148. [doi: 10.7326/M21-2972] [Medline: 34748380]
68. Li W, Ornstein KA, Li Y, Liu B. Barriers to learning a new technology to go online among older adults during the COVID-19 pandemic. J Am Geriatr Soc 2021 Nov;69(11):3051-3057. [doi: 10.1111/jgs.17433] [Medline: 34409589]
69. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: A systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
70. Choi NG, DiNitto DM, Marti CN, Choi BY. Telehealth use among older adults during COVID-19: associations with sociodemographic and health characteristics, technology device ownership, and technology learning. J Appl Gerontol 2022 Mar;41(3):600-609. [doi: 10.1177/07334648211047347] [Medline: 34608821]
71. Harju A, Neufeld J. Telehealth utilization during the COVID-19 pandemic: a preliminary selective review. Telemed Rep 2022;3(1):38-47. [doi: 10.1089/tmr.2021.0040] [Medline: 35720447]
72. Zubair Elahi M, Liang G, Jawad Malik M, Dilawar S, Ilyas B. Fear of Covid-19 and intentions towards adopting E-Health services: exploring the technology acceptance model in the scenario of pandemic. International Journal of Business, Economics and Management 2021;8(4):270-291 [FREE Full text] [doi: 10.18488/journal.62.2021.84.270.291]
73. Jacobs JM, Rottenberg Y, Cohen A, Stessman J. Physical activity and health service utilization among older people. J Am Med Dir Assoc 2013 Feb;14(2):125-129. [doi: 10.1016/j.jamda.2012.10.023] [Medline: 23218744]
74. Kang SW, Xiang X. Physical activity and health services utilization and costs among U.S. adults. Prev Med 2017 Mar;96:101-105. [doi: 10.1016/j.ypmed.2016.12.043]
75. Gobbi S, Płomecka MB, Ashraf Z, et al. Worsening of preexisting psychiatric conditions during the COVID-19 pandemic. Front Psychiatry 2020;11:581426. [doi: 10.3389/fpsyt.2020.581426] [Medline: 33391049]
76. Robbins R, DiClemente RJ, Baig N, Johnson A, Chou A, Van den Bulck J. Digital communications technology use and feelings of anxiety, depression, and loneliness among older adults during the COVID-19 pandemic. J Appl Gerontol 2023 Sep;42(9):1911-1920. [doi: 10.1177/07334648231169086] [Medline: 37041750]
77. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024 Feb 21;10:e48685. [doi: 10.2196/48685] [Medline: 38381497]
78. Grant N, Wardle J, Steptoe A. The relationship between life satisfaction and health behavior: A cross-cultural analysis of young adults. Int J Behav Med 2009;16(3):259-268. [doi: 10.1007/s12529-009-9032-x] [Medline: 19319695]
79. Martín-María N, Caballero FF, Moreno-Agostino D, et al. Relationship between subjective well-being and healthy lifestyle behaviours in older adults: a longitudinal study. Aging Ment Health 2020 Apr 2;24(4):611-619. [doi: 10.1080/13607863.2018.1548567]
80. Bai A, Li H, Huang Y, et al. A survey of overall life satisfaction and its association with breast diseases in Chinese women. Cancer Med 2016 Jan;5(1):111-119. [doi: 10.1002/cam4.565] [Medline: 26640035]
81. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. NPJ Digit Med 2020;3(1):94. [doi: 10.1038/s41746-020-0306-7] [Medline: 32685700]
82. Raja M, Bjerkan J, Kymre IG, Galvin KT, Uhrenfeldt L. Telehealth and digital developments in society that persons 75 years and older in European countries have been part of: a scoping review. BMC Health Serv Res 2021 Oct 26;21(1):1157. [doi: 10.1186/s12913-021-07154-0] [Medline: 34696789]
83. Schiel S, Marcos J, Schulz S, Dickmann C, Alterssurvey F. Methodenbericht deutscher alterssurvey (DEAS): durchführung der 7. erhebungswelle 2020/2021 [article in german]. : INFAS Institut für angewandte Sozialwissenschaft GmbH; 2021 URL: https://www.ssoar.info/ssoar/bitstream/handle/document/79743/ssoar-2021-schiel_et_al-Methodenbericht_Deutscher_Alterssurvey_DEAS_Durchfuhrung.pdf?sequence=1&isAllowed=y&lnkname=ssoar-2021-schiel_et_al-Methodenbericht_Deutscher_Alterssurvey_DEAS_Durchfuhrung.pdf [accessed 2025-05-23]

Abbreviations

CES-D: Center of Epidemiologic Studies - Depression scale
DEAS: Deutsches Alterssurvey (German Ageing Survey)
ISCED: International Standard Classification of Education
OR: odds ratio
SHARE: Survey of Health, Ageing and Retirement in Europe
WHO: World Health Organization

Edited by G Seçkin; submitted 07.05.24; peer-reviewed by D Singh, L Mao, W Li; revised version received 06.02.25; accepted 23.02.25; published 26.05.25.

Please cite as:
Neumann A, König HH, Hajek A
Determinants of Having Online Health Consultations During the COVID-19 Pandemic Among Middle-Aged and Older Adults in Germany: Representative Longitudinal Survey Study
JMIR Aging 2025;8:e60311
URL: https://aging.jmir.org/2025/1/e60311
doi:10.2196/60311

Analyzing Disparity in Geographical Accessibility to Home Medical Care Using a Claims Database and Geographical Information System: Simulation Study

Yasuhiro Morii[1], PhD; Yasuhiro Nakanishi[2], PhD; Yuichi Nishioka[3], PhD; Yukio Tsugihashi[3,4], PhD; Tatsuya Noda[3], PhD; Tomoya Myojin[3,5], PhD; Tomoaki Imamura[3], PhD; Manabu Akahane[2], PhD

[1]Center for Outcomes Research and Economic Evaluation for Health, National Institute of Public Health, Wako, Japan

[2]Department of Health and Welfare Services, National Institute of Public Health, 2-3-6 Minami, Wako, Japan

[3]Department of Public Health, Health Management and Policy, Nara Medical University, Kashihara, Japan

[4]Medical home care center, Tenri Hospital Shirakawa Branch, Tenri, Japan

[5]Department of Community Health and Preventive Medicine, Hamamatsu University School of Medicine, Hamamatsu, Japan

Corresponding Author:
Manabu Akahane, PhD
Department of Health and Welfare Services, National Institute of Public Health, 2-3-6 Minami, Wako, Japan

Abstract

Background: The demand for home medical care services has increased in aging societies. Therefore, allocating health care resources optimally to meet the needs of each community is essential. Geographical accessibility is an important factor affecting access to home medical care services; however, little research has been conducted on regional disparities in geographical accessibility.

Objective: This study aims to analyze the regional disparities in geographical accessibility to home medical care services using the Kokuho database (KDB), a comprehensive medical claims database for a prefecture in Japan.

Methods: This study included 39 municipalities in Nara Prefecture, Japan. Using a geographical information system, accessibility to home medical care services, that is, travel distance and time from hospitals and clinics to hypothetical patients, was analyzed in two scenarios: (1) an ideal scenario, where we assumed that all hospitals or clinics in Nara Prefecture provided those services and (2) an actual scenario, where hospitals or clinics in Nara Prefecture that actually provided home medical care services, identified from KDB data analysis, were used in the analysis. Hypothetical patients were randomly distributed on the geographical information system in accordance with the usage rates of home medical care services and with the distributions of the population aged ≥75 years. The usage rate by municipalities was aggregated from the analysis of KDB data of Nara Prefecture in FY2019.

Results: The median travel distance was longer than 16 km, the reference limit value specified in the Japanese fee table, and the median travel time exceeded 30 min in certain rural municipalities in the southern part of Nara Prefecture, in the actual scenario, whereas the travel distance and time were improved in the ideal scenario. The differences in travel time between the ideal and actual scenarios were the largest in the depopulated municipalities in the southern part, such as Totsukawa (32.6 vs 5.8 min), Kawakami (30.1 vs 11.8 min), Kurotaki (21.3 vs 5.2 min), and Kamikitayama (20.7 vs 3.5 min). The usage rates were also lower in rural municipalities in the southern part.

Conclusions: The results revealed that geographical accessibility was lower in depopulated municipalities in the southern part, and the disparity could be partly solved in the ideal scenario, especially in that area, highlighting the necessity of increasing supply in the southern areas. KDB is a comprehensive database that includes medical claims information for home medical care patients and details of the provision of medical institutions, enabling geographical analysis that reflects actual health care usage.

(JMIR Aging 2025;8:e70040) doi:10.2196/70040

KEYWORDS

home medical care; geographical accessibility; geographical information system; medical claims database; simulation

Introduction

Most limited-income countries have aging societies, including Japan, which is the most aging country in the world with an aging rate (rate of population aged ≥65 y) of 28.1% in 2018, followed by Italy (23.3%), Portugal (21.9%), and Germany (21.7%) [1]. Medical demand is estimated to increase with demographic changes [2]. The Japanese government has been promoting a structure in which home medical care and long-term care services are provided comprehensively to older adults,

ensuring they live their lives in the areas they are used to in an aging society [3]. Therefore, increased promotion of the use of home medical care services is of great societal importance.

The demand for home medical care services has also increased in recent years, with a peak at around 2040 [3]. Each prefecture in Japan is responsible for creating a Regional Medical Plan to ensure that health care services are effectively provided to meet the needs of the local population [4]. The Regional Medical Plan includes a plan to provide home medical care services, and each prefecture must specify the problems faced and measures to resolve them. Prefectures must consider optimal allocations and make the most of their limited resources to meet the needs.

Geographical accessibility is among the factors that could affect the provision of home medical services. The relationship between geographical factors and the usage of medical services has been previously studied [5-7]. For example, Schwarz et al [5] analyzed the relationship between the distance from hospitals to patients and the psychiatric treatments that patients received. Herein, people living in peripheral communities were mainly treated in inpatient settings. In contrast, those living in residential areas were more likely to receive intensive psychiatric home treatment. Similar to their study findings, geographical accessibility is important for home medical care services since health care providers need to travel to patients' homes.

Kuwayama et al [8] conducted a literature review of studies on home care nursing in Japan using a geographical information system (GIS) and discovered a few related studies [9,10]. In those studies, Naruse et al [7] performed a multilevel logistic analysis of the association between the use of home care nursing and the proportion of older adults reachable by facilities providing home-visit nursing services (calculated by buffer analysis using GIS). They found that people living in less reachable municipalities were less likely to use these services. However, few studies have been conducted on the geographical accessibility of home-based medical care services. Moreover, these previous studies were limited to visualizing the current situation.

Extensive databases containing medical claims and national health insurance information, such as the National Database of Health Insurance Claims and Specific Health Checkups of Japan (NDB) and National Health Insurance Database (Kokuho database: KDB), have recently become essential tools for medical and health care policy research. These real-world data sources provide valuable insights into health care usage, treatment patterns, and outcomes. The main advantages of using such databases include the ability to study targeted populations and understand the actual clinical practice comprehensively. Using such databases, a geographical analysis that reflects the actual health care usage and trends will be possible.

Several studies have conducted geographical analysis using a large real-world database. Cheng et al [11] analyzed the prescribing patterns of cardiovascular drugs across townships in Taiwan using a national claim database covering 97.5% of the total population. Kwak et al [12] analyzed the relationship between the incidence of obstetric complications for pregnant women and access time to delivery units in South Korea, using the national claim database.

However, there are few previous studies analyzing geographical accessibility to home medical care services using an extensive database.

Therefore, this study aimed to develop a methodology for analyzing regional disparities in geographical accessibility to home medical care services using the KDB database to help consider achieving the provision of more efficient and equitable service. Given the global trend of aging, the findings of this study in Japan may contribute to improving the efficiency and equity of service delivery and policymaking of home medical care services in other countries facing similar demographic challenges.

Methods

Subjects and Scheme

This study was conducted in the 39 municipalities in Nara Prefecture, Japan, as a model region. Nara Prefecture is located in the center of Japan's main island of Honshu (Figure 1A) and has both urban and rural (mountainous) areas. Nara Prefecture has a population of 1,286,651 (approximately 1/100th of the national total), a land area of 3691 km^2 (approximately 1/100th of the national total) [13,14], and 32% (0.4/1.25) of adults aged ≥65 years (compared to 36.4/125.52, 29% nationally; Table 1). The characteristics of municipalities, such as population and aging rate, are shown in Table 1. Figures 1B and C illustrate the geographical locations of the 5 secondary medical areas and 39 municipalities. Nanwa secondary area, in the southern part, is the most rural area with mountains. Since there seems to be a geographical difference within Nara Prefecture, it would be suitable as a model region for geographical analysis. Secondary medical areas provide general inpatient medical services, whereas municipalities or smaller regional units provide more basic medical services.

Figure 1. The geographical locations of Nara Prefecture in Japan, secondary medical areas in Nara, and municipalities in Nara show the geographical location of (A) Nara Prefecture in Japan and also illustrate the distributions of hospitals or clinics (B) in the actual scenario and (C) in the ideal scenario.

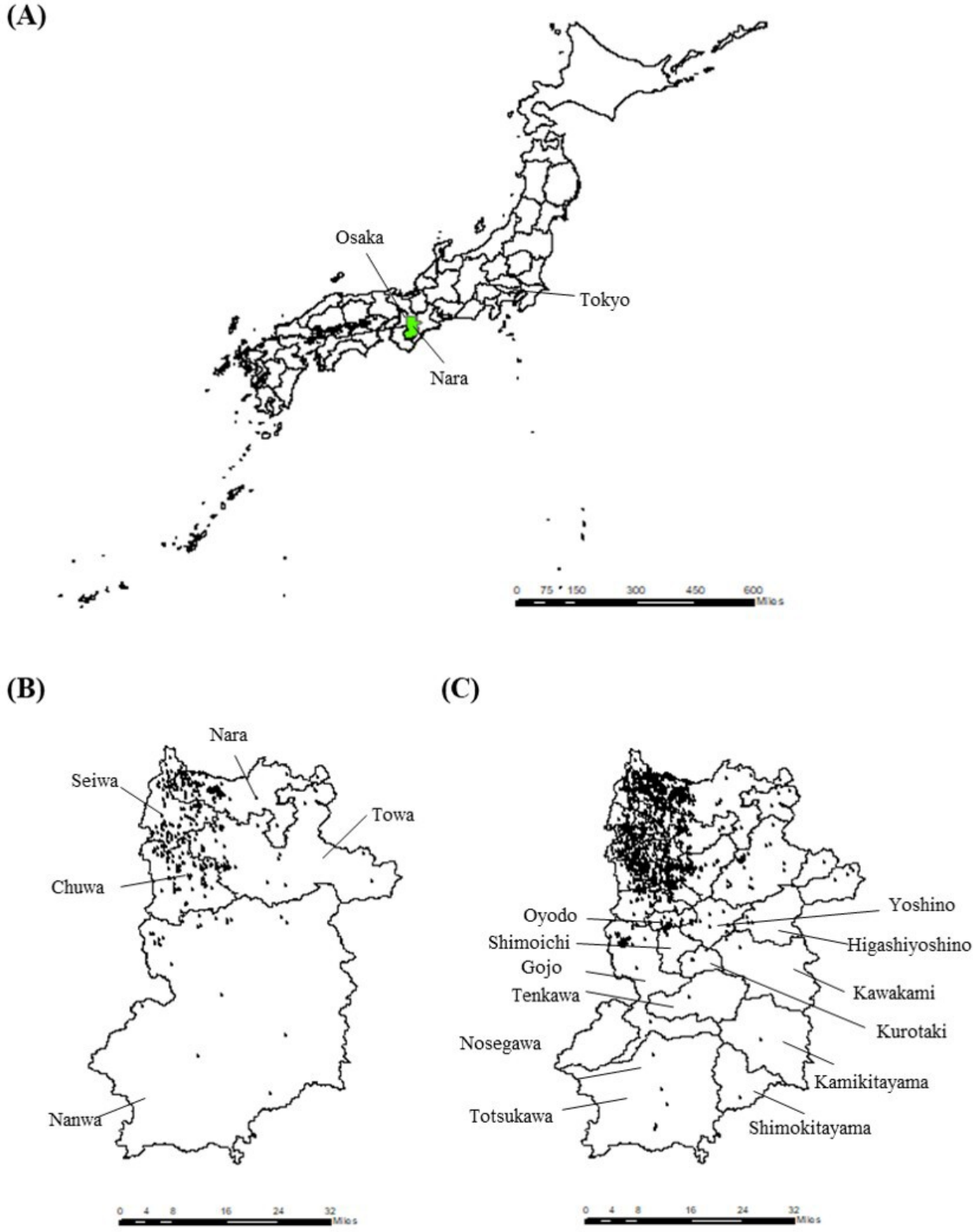


Table . Characteristics of each municipality and the result of usage rate.

Municipality	Population [14]	Aging rate (%) [14]	Land area (km^2) [13]	Population density (/km^2)	Number of claims	Usage rate (%)
Nara	347,947	32.5	276.9	1256.40	34,335	6.8
Kashihara	118,344	30	39.6	2991.50	5370	3.9
Ikoma	114,508	30	53.2	2154.40	7755	5.8
Yamatokoriyama	80,760	34.4	42.7	1891.80	5135	4.6
Kashiba	77,212	24.8	24.3	3182.70	3931	5.4
Tenri	60,717	27.9	86.4	702.6	4160	5.8
Yamatotakada	59,797	32.9	16.5	3628.50	3838	4.4
Sakurai	52,975	33.1	98.9	535.6	3812	5.4
Katsuragi	37,199	28.1	33.7	1103.20	1424	3.4
Koryo	33,910	27.3	16.3	2080.40	1154	3.6
Tawaramoto	30,840	32.6	21.1	1462.30	1497	4.9
Ikaruga	27,408	30.7	14.3	1920.70	2108	6.0
Uda	25,732	44.5	247.5	104	2103	4.7
Gojo	25,487	42.4	292	87.3	1769	4.2
Oji	23,667	28.9	7	3376.20	1518	5.4
Sango	22,688	33	8.8	2581.10	2018	7.2
Gose	22,259	43.9	60.6	367.4	1096	2.6
Kanmaki	20,815	37	6.1	3390.10	1586	5.3
Heguri	17,545	39.9	23.9	734.1	2484	7.9
Kawai	16,314	40.2	8.2	1982.30	1523	5.3
Oyodo	15,530	37.5	38.1	407.6	381	2.2
Kawanishi	7716	36	5.9	1301.20	457	3.6
Ando	6930	37.2	4.3	1607.90	439	4.9
Takatori	6235	43.6	25.8	241.8	274	2.5
Miyake	6097	37.2	4.1	1501.70	457	4.7
Yoshino	5457	54	95.7	57.1	407	2.7
Asuka	4767	43.2	24.1	197.8	238	3.1
Shimoichi	4356	49.7	62	70.3	154	2.0
Yamazoe	2909	51.6	66.5	43.7	289	4.4
Totsukawa	2695	43.4	672.4	4	254	3.6
Higashiyoshino	1337	60.5	131.7	10.2	224	5.1
Mitsue	1298	61.2	79.6	16.3	50	1.6
Soni	1175	53.5	47.8	24.6	140	5.0
Tenkawa	1048	53.3	175.7	6	60	1.5
Kawakami	1039	56.1	269.3	3.9	112	2.7
Shimokitayama	690	47.9	133.4	5.2	39	2.1
Kurotaki	536	54	47.7	11.2	24	0.9
Kamikitayama	384	49.6	274.2	1.4	37	2.6
Nosegawa	328	51.6	154.9	2.1	62	10.5
Total	1,286,651	32.7	3690.9	348.6	92,714	5.3

For each municipality, geographical accessibility to home medical care services was analyzed using GIS (ArcGIS Desktop version 10.8.1; ESRI Japan Corporation) in the following two scenarios: (1) actual scenario: geographical accessibility to patients from hospitals and clinics in Nara Prefecture that actually provided home medical care services, and (2) ideal scenario: geographical accessibility to patients from all hospitals or clinics in Nara Prefecture was used. The latter simulates geographical accessibility in an ideal scenario in which the number of hospitals or clinics that provide home medical care services would increase. These scenarios were different only in terms of the number of hospitals and clinics used in the analysis, as described hereafter. Hypothetical patients were generated on a GIS by using the methodology described in a subsequent section "Geographical Analysis."

Data

For the actual scenario, hospitals and clinics that provided home medical care services were defined using data from the KDB database of Nara Prefecture (Nara KDB) in FY2019. This is a prefecture-wide database of medical and long-term care insurance claims. To be defined as hospitals and clinics that actually provided home medical care services, hospitals and clinics were required to have at least one claim with home medical care service fees (medical service code: 114001110, 114030310, 114042110, 114042210, 114042810, 114046310, 114027710, and 114027810 in the national fee schedule) for patients aged ≥75 years in Nara KDB in FY2019. Consequently, 346 hospitals and clinics were identified in the actual scenario.

The ideal scenario used all the hospitals and clinics in Nara prefecture, that is, those providing or not providing home medical care services. The geographical data on the 1308 hospitals and clinics were obtained from the Geospatial Information Authority of Japan [15]. Dental clinics were excluded from the study. The distribution of hospitals and clinics in the 2 scenarios is shown in Figure 1.

Population data at each 500 m^2 mesh (ESRI Japan Corporation) were used to distribute the locations of the hypothetical patients.

Geographical Analysis

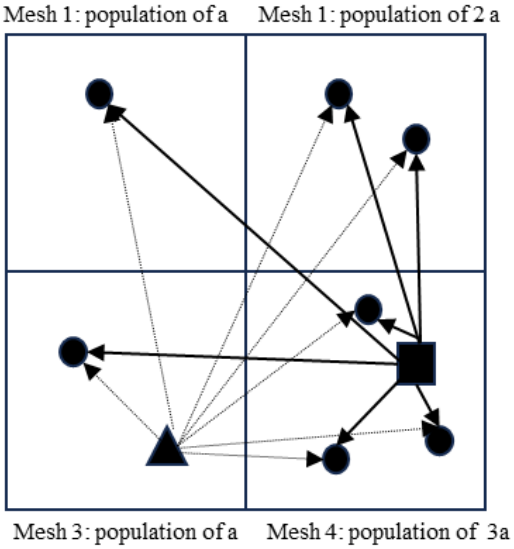
First, hypothetical patients were generated on GIS. The number of patients in each municipality was that of patients aged ≥75 years who used home medical care services in each municipality, identified from the Nara KDB data analysis. The usage of home medical care was defined in the same manner as identifying the hospitals and clinics in the actual scenario. That number of patients was randomly allocated to 500 m^2 meshes on GIS, reflecting the distribution of the population aged ≥75 years based on the national census 2020. In other words, the hypothetical patients were allocated to each mesh in a manner that the number of patients in each mesh is proportional to the mesh population aged ≥75 years using R version 4.1.2 (The R Foundation) [16,17]. This methodology enables the simulation of accessibility, reflecting the actual geographical distribution of potential patients. Thereafter, based on the patient allocations, hypothetical patient points were created on randomly determined points within the allocated meshes on ArcGIS using the "Create Random Points" function.

Subsequently, travel time from the patients to the hospitals or clinics was analyzed using the "Find Closest Facilities" function of ArcGIS for the 2 scenarios, respectively. We assumed that the travel to the patients was from the closest hospital or clinic. However, the closest hospitals or clinics can be located outside the municipality where the patients live.

The median and 25 - 75 percentiles of travel distance and time were aggregated for the patients in each municipality. Moreover, the rates of patients for whom travel distances were over 16 km were calculated since, in the Japanese universal fee table, the travel distance to provide home medical care services needs to be 16 km or less, unless an exceptional reason is present [18].

The usage rate was calculated by dividing the number of patients aged ≥75 years using home medical care services by the total population aged ≥75 years. The entire process of geographical analysis was repeated 10 times to consider the uncertainty of patient locations in each trial. The scheme of geographical analysis is visualized in Figure 2.

Figure 2. Scheme of the geographical analysis illustrating the scheme of the geographical analysis. GIS: geographical information system.



1. The hospitals or clinics defined to provide home medical care services (■) were identified from claims database analysis
2. All hospitals and clinics (both ▲ and ■) were identified from an open source
3. Locations of hypothetical patients were randomly decided in a manner to reflect actual distributions of population over 75 or more and usage rates for each municipality (using R 4.1.2) (eg. in a mesh with "2a" population likely has twice more patients than that with "a" population on average)
4. The hypothetical patients were mapped on GIS
5. Travel distance and time were analyzed using GIS
   - in ideal scenario: both ▲ and ■ were used
   - in actual scenario: only ■ was used (solid arrows)

Ethical Considerations

This study was approved by the ethics committee of the National Institute of Public Health, Japan (NIPH-IBRA number 12324 - 2). The requirement for informed consent was waived because all data were anonymized.

Results

The number of claims and usage rates are presented in Table 1. The distributions of hospitals or clinics used in each scenario were visualized in Figure 1. The travel time and distance results are shown in Figures 3-6 and Multimedia Appendices 1 and 2. In the actual scenario, the median travel time for all the patients in Nara Prefecture was 3.0 (IQR 2.0 - 4.3) minutes, and the travel distance was 0.7 (IQR 0.5 - 1.2) km. In the ideal scenario, the median travel time for all patients was 2.0 (IQR 1.2 - 2.9) minutes, and the travel distance was 0.5 (IQR 0.3 - 0.9) km.

Figure 3. The median travel time in the two scenarios. This figure illustrates the median travel time in (A) the ideal scenario and (B) the actual scenario.

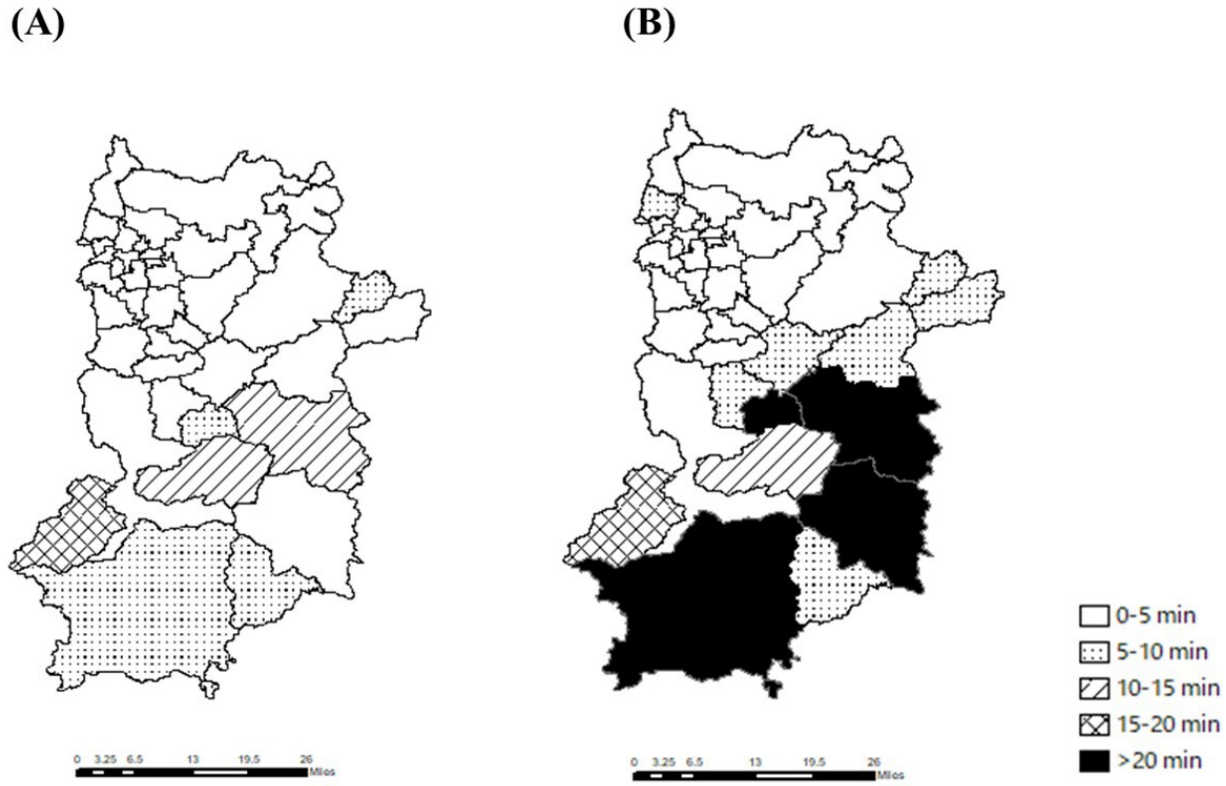


Figure 4. The rates of patients for whom the travel time was longer than 30 minutes. This figure illustrates the rates of patients for whom the travel time was longer than 30 minutes in (A) the ideal scenario and (B) the actual scenario.

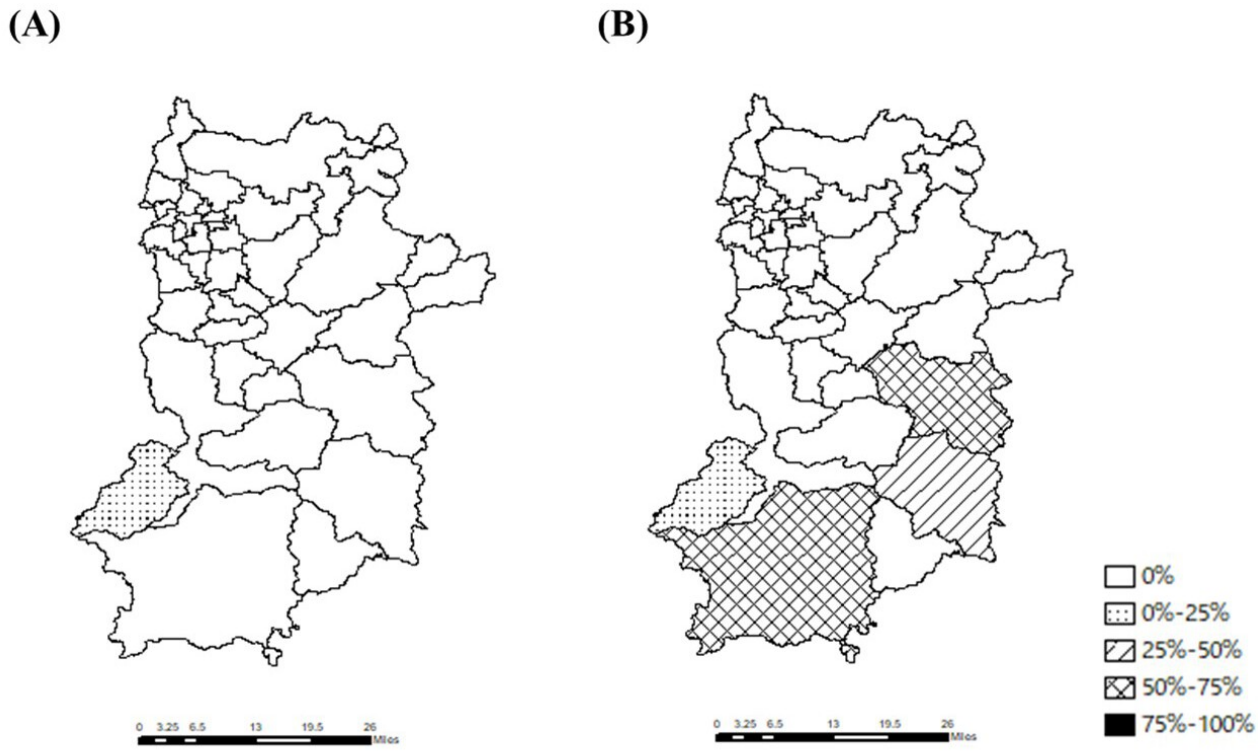


Figure 5. The median travel distance for each municipality in the two scenarios illustrates the median travel distance in (A) the ideal scenario and (B) the actual scenario.

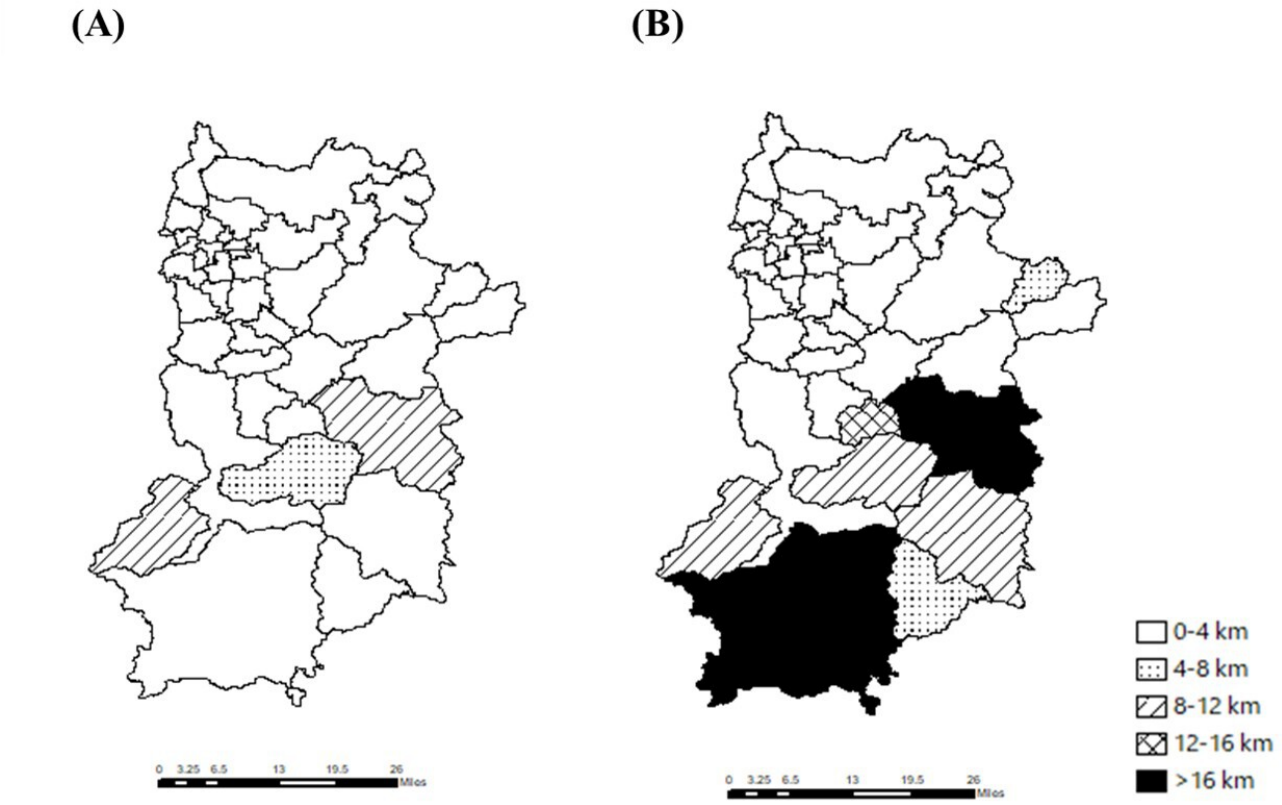
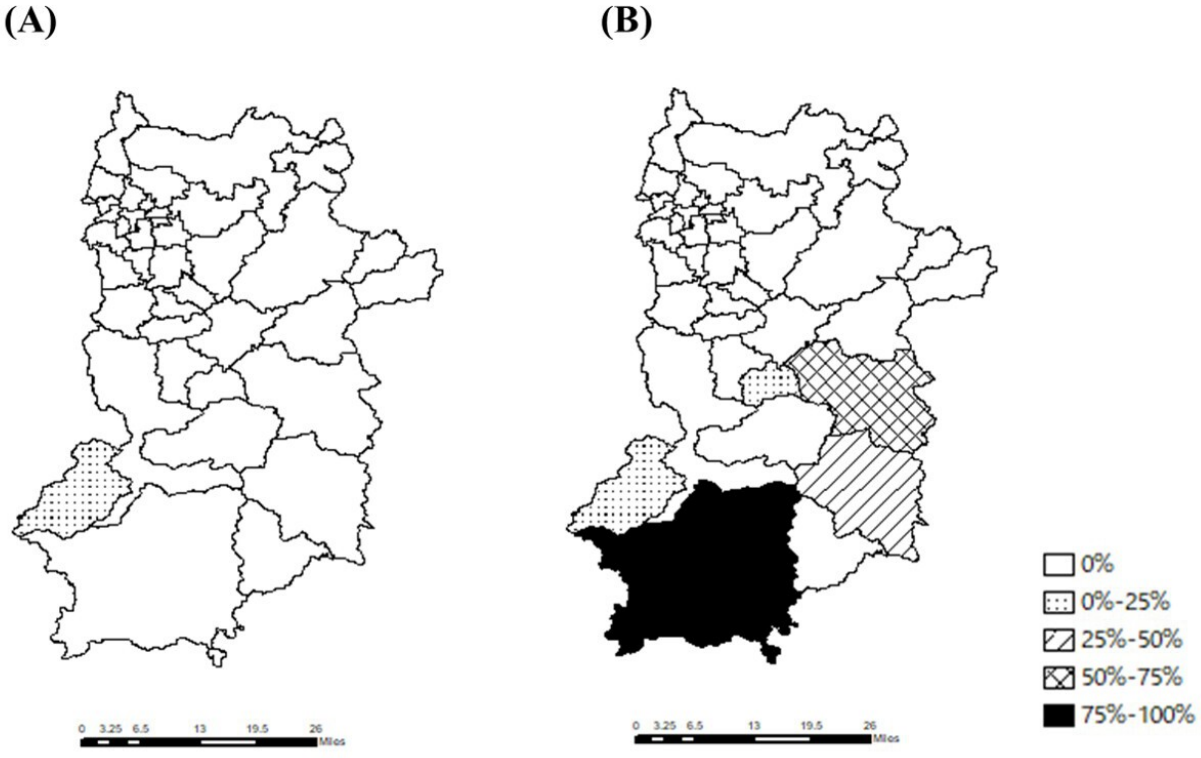


Figure 6. The rates of patients for whom the travel distance was longer than 16 km. This illustrates the rates of patients for whom the travel distance was longer than 16 km in (A) the ideal scenario and (B) the actual scenario.



The median travel times for each municipality in the 2 scenarios are shown in Figures 3 and 4. In the ideal scenario, the median travel time was less than 15 minutes, except for Nosegawa (18.1, IQR 2.2 - 25.5 min), while in the actual scenario, the median travel time was longer in the municipalities in the Nanwa area, the southern part (in Totsukawa (32.6, IQR 21.2 - 40.5 min), Kawakami (30.1, IQR 16.5 - 32.4 min), Kurotaki (21.3, IQR 20.2 - 24.1 min), and Kamikitayama (20.7, IQR 19.7 - 26.1 min; Figure 3).

When comparing the 2 scenarios, the differences in travel time between the scenarios were the largest in Totsukawa (26.8 min), Kawakami (18.3 min), Kamikitayama (17.2 min), and Kurotaki (16.1 min). All of these 4 municipalities are located in a depopulated secondary medical area of Nanwa.

Figure 4 shows the number of patients for whom the travel time was longer than 30 minutes. In the actual scenario, the rates were highest in Totsukawa (58.0%), Kawakami (50.0%), Kamikitayama (30.0%), and Nosegawa (25.0%).

The median travel distances for each municipality in the 2 scenarios are shown in Figure 5. In the ideal scenario, the median travel distances were shorter than 12 km for all municipalities. However, in the actual scenario, the median travel distances were longer than 16 km in Totsukawa (28.4, IQR 16.7 - 32.3 km) and Kawakami (17.2, IQR 10.3 - 19.3 km). Kurotaki, Nosegawa, and Kamikitayama showed similar findings (13.9, IQR 12.6 - 14.4; 11.5, IQR 1.3 - 16.1; and 10.8, IQR 10.3 - 16.4 km, respectively). The differences in travel distance between the scenarios were the largest in Totuskawa (24.5 km), Kurotaki (11.8 km), Kamikitayama (9.3 km), and Kawakami (8.6 km).

Figure 6 shows the number of patients whose travel distance was >16 km. In the actual scenario, the rates were 77.6% for Totsukawa, 56.2% for Kawakami, 30.0% for Kamikitayama, 25% for Nosegawa, 20% for Kurotaki, and 0% for the other municipalities. In the ideal scenario, the rate was zero, except in Nosegawa (25.4%).

The usage rates by municipalities are shown in Table 1. This table shows the usage rates of home medical care services in each municipality. As can be seen, the usage rates tended to be lower in the municipalities in the Nanwa area (colored black), except for Nosegawa, which had the highest rate.

Discussion

Principal Findings

In this study, an accessibility analysis of home medical care services in Nara Prefecture was conducted as a model region, using the KDB data to investigate the equitable distribution of home medical care services. Using the KDB, we accurately identified medical institutions that provide home medical care, estimated the usage rates of home medical care services for each municipality, and demonstrated the feasibility of analyzing geographical accessibility to home medical care services that reflect actual health care usage and clinical trends.

The findings of this study revealed that in certain municipalities within the Nanwa area, a rural secondary medical area with a declining population, the travel distance to a health care patient's home exceeded 16 km, and the travel time exceeded 30 minutes in some cases. These results highlight disparities among municipalities in terms of geographic accessibility to home medical care services. Its supply in the Nanwa area needs to be increased from the perspective of geographic accessibility.

In addition, the usage rates of these services were lower in rural municipalities in the Nanwa area. To promote the use of home medical care services, improving accessibility in these municipalities is necessary. Geographical accessibility—that is, the distance from service providers to patients—could affect service usage. Naruse et al [7] reported that travel distance can affect the usage of home nursing care services. Although geographical accessibility would also likely affect the usage of home medical care services, the extent of the effect has not been clear since this study did not quantify the effect by controlling other factors, such as socioeconomic status, care quality, and regional characteristics. Further studies with additional information beyond medical claim data are required to clarify this point.

In recent years, the government has promoted the use of home medical care to meet the increasing medical demands within each area, enabling local people to live their lives in areas they have always lived in [3]. Our results showed that a disparity occurred in geographical accessibility to home medical care based on the current practice in the actual scenario; however, in the ideal scenario, no municipality was observed in which the travel distance to any patient was longer than 16 km, except for Nosegawa. Geographical accessibility is much better in this scenario. For example, the average travel time was less than 15 minutes, except for Nosegawa. These results indicate that the disparity in terms of geographical accessibility would have been mostly resolved if all hospitals and clinics had provided these services. Furthermore, our results indicated that the potential benefit would be the largest in the Nanwa area, the most depopulated area. Therefore, if the number of hospitals or clinics providing these services is increased by a policy (eg, through dispatching health care professionals or funding), this will be effective in ensuring sufficient geographical accessibility for patients in the Nanwa area (specifically, Kawakami, Totsukawa, Kurotaki, and Kamikitayama). That is especially necessary in Kawakami and Kurotaki, since no hospitals or clinics were providing services, and the usage rates were low in these municipalities. However, it is neither feasible nor efficient for all hospitals and clinics to provide services because of limited resources. Further analysis using GIS could clarify how residents' geographical accessibility will change when a specific hospital or clinic starts to provide home medical care services.

Notably, while the travel distance was longer in Nosegawa than in the other municipalities in both scenarios, the usage rate was the highest. The high usage rate could have been achieved by factors other than geographical accessibility. Further study, including factors other than geographical accessibility, is required to clarify the reason.

Our study is among the few that have analyzed geographical accessibility to home medical care services using GIS and KDB. The Nara KDB used in this study is an integrated database of

medical and long-term care insurance claim information covering all residents aged 75 years and older, which matches the primary population of home medical care patients. This methodology enabled us to conduct a geographical analysis that reflected actual health care usage. Furthermore, our results are important when considering the optimal allocation of medical resources related to home medical care services and equal accessibility to these services. Japan is the most aging country in the world [1]. Home medical care services will be more important in other countries, and ensuring accessibility to these services will be of greater importance in the future. Geographical analyses using nationwide health care claims databases have been conducted in other countries for health care fields, such as drug prescription and pregnancy outcomes [11,12]. Likewise, our methodology can be applied to other countries for home medical care services, as those services will be more important as the aging of the population progresses.

Limitations

The KDB is a comprehensive database that includes medical claims information for home medical care patients and details the provision of medical institutions, enabling a geographical analysis that reflects actual health care usage. However, this study had some limitations. First, it did not consider the capacity of each hospital or clinic because it focused on geographical accessibility. Further studies are required to consider its capacity to be more informative for optimal resource allocation. Second, a hospital or clinic was regarded as providing home medical care services if at least one claim regarding services from a hospital or clinic was identified. Some hospitals and clinics included in the actual scenario might not routinely provide home medical care services. Third, the disparity in geographical accessibility within each municipality is beyond the scope of this study. Such disparities may exist particularly in municipalities with lower geographical accessibility. Fourth, the KDB was used to analyze home medical care patients aged ≥75 years; therefore, home health care patients aged <75 years were not included. In the future, using the NDB or KDB and applying the developed methods, the analysis could be expanded to include pediatric home medical care patients and those under the age of 75 years. Finally, this study only used KDB data from FY 2019 to conduct geographical analysis while excluding the impact of the COVID-19 pandemic. The usage of home medical care has been increasing over the years, and the usage patterns could also change over the years. By applying the methodology established in this study, a nationwide analysis using the most recent data, free from the influence of the COVID-19 pandemic, can provide valuable insights into the future of home medical care.

Conclusions

This study conducted a geographical accessibility analysis of home medical care services using Nara Prefecture as a model region with GIS and KDB. The results revealed that geographical accessibility was lower in depopulated municipalities in the southern part, such as Totsukawa, Kawakami, Kamikitayama, and Nosegawa, which have lower accessibility than urban areas. However, the results also showed that the disparity could be partly solved in the ideal scenario. The differences in travel time between the scenarios were especially larger in the depopulated areas in the southern part, such as Totsukawa (32.6 vs 5.8 min), Kawakami (30.1 vs 11.8 min), Kurotaki (21.3 vs 5.2 min), and Kamikitayama (20.7 vs 3.5 min). These results highlight the necessity of increasing supply in the southern areas.

Acknowledgments

This study was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare, Japan.

Data Availability

The datasets used in the current study are not publicly available owing to the publication policy of the Nara Prefecture but are available from the corresponding author on reasonable request subject to approval from the Nara Prefecture.

Authors' Contributions

YM, YN, and MA designed the research plan. YT, TN, TM, and TI contributed significantly to the development of the research. YM performed the geographical analysis and drafted the manuscript. YN performed the claims database analysis. AM led the research project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Median and 25-75 percentiles of travel distance in the ideal and actual scenarios (km).
[DOCX File, 24 KB - aging_v8i1e70040_app1.docx]

Multimedia Appendix 2
Median and 25-75 percentiles of travel time in the ideal and actual scenarios (min).

[DOCX File, 23 KB - aging_v8i1e70040_app2.docx]

References

1. International comparison of elderly population (in japanese). Statistics Bureau of Japan. URL: https://www.stat.go.jp/data/topics/topi1135.html [accessed 2025-07-01]
2. Constructing the structure of providing home medical care services (in japanese). The Ministry of Health, Labour and Welfare. 2023. URL: https://www.mhlw.go.jp/content/10800000/001146149.pdf [accessed 2024-09-07]
3. Home medical care services -toward 8th regional medical plan (in japanese). The Ministry of Health, Labour and Welfare. 2021. URL: https://www.mhlw.go.jp/content/10800000/000911334.pdf [accessed 2024-09-07]
4. Regional Medical Plan (In Japanese). The Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryou/iryou_keikaku/index.html [accessed 2024-09-07]
5. Schwarz J, Hemmerling J, Kabisch N, et al. Equal access to outreach mental health care? Exploring how the place of residence influences the use of intensive home treatment in a rural catchment area in Germany. BMC Psychiatry 2022 Dec 26;22(1):826. [doi: 10.1186/s12888-022-04477-y] [Medline: 36572855]
6. Wang Y, Zhang Q, Spatz ES, et al. Persistent geographic variations in availability and quality of nursing home care in the United States: 1996 to 2016. BMC Geriatr 2019 Apr 11;19(1):103. [doi: 10.1186/s12877-019-1117-z] [Medline: 30975076]
7. Naruse T, Matsumoto H, Fujisaki-Sakai M, Nagata S. Measurement of special access to home visit nursing services among Japanese disabled elderly people: using GIS and claim data. BMC Health Serv Res 2017 May 30;17(1):377. [doi: 10.1186/s12913-017-2322-0] [Medline: 28558677]
8. Kuwayama T, Hamabata K, Kamesaki T, Koike S, Kotani K. Research on home care nursing in Japan using geographic information systems: a literature review. Jpn Clin Med 2018;9:1179670718814539. [doi: 10.1177/1179670718814539] [Medline: 30515029]
9. Shimizu J, Osanai S. Feasibility of home visit by nurses with higher palliative care specialty with visiting nurses in Japan—analysis of nationwide distribution and geographical associations [Abstract in English]. J Jpn Health Sci 2014;16:177-183.
10. Iizaka S, Sanada H. Distribution of Tokyo treatment centers with wound, ostomy, and continence nurses: a spatial analysis of hospitals and municipalities using a geographic information system. Jpn J PU 2016;18:118-126.
11. Cheng CL, Chen YC, Liu TM, Yang YHK. Using spatial analysis to demonstrate the heterogeneity of the cardiovascular drug-prescribing pattern in Taiwan. BMC Public Health 2011 May 24;11:380. [doi: 10.1186/1471-2458-11-380] [Medline: 21609462]
12. Kwak MY, Lee SM, Kim HJ, et al. How far is too far? A nationwide cross-sectional study for establishing optimal hospital access time for Korean pregnant women. BMJ Open 2019 Sep 20;9(9):e031882. [doi: 10.1136/bmjopen-2019-031882] [Medline: 31542767]
13. Geospatial information authority of japan. Land area (In Japanese). URL: https://www.gsi.go.jp/KOKUJYOHO/MENCHO/backnumber/GSI-menseki20250401.pdf [accessed 2024-09-19]
14. Nara prefectural government estimated population. Estimated population. 2024. URL: https://www.pref.nara.jp/6265.htm [accessed 2024-09-07]
15. Geospatial information authority of Japan. Digital Land Information Download Website (In Japanese). URL: https://nlftp.mlit.go.jp/ksj/index.html [accessed 2024-09-07]
16. Morii Y, Osanai T, Fujiwara K, et al. Analyzing Cost-Effectiveness of Allocating Neurointerventionist for Drive and Retrieve System for Patients with Acute Ischemic Stroke. J Stroke Cerebrovasc Dis 2021 Aug;30(8):105843. [doi: 10.1016/j.jstrokecerebrovasdis.2021.105843] [Medline: 34000607]
17. Morii Y, Osanai T, Ishikawa T, et al. Cost Effectiveness of Drive and Retrieve System in Hokkaido for Acute Ischemic Stroke Patient Treatment Using Geographic Information System. J Stroke Cerebrovasc Dis 2019 Aug;28(8):2292-2301. [doi: 10.1016/j.jstrokecerebrovasdis.2019.05.020] [Medline: 31200963]
18. A requirement on distance for home medical care services. The Ministry of Health, Labour and Welfare. 2023. URL: https://www.mhlw.go.jp/stf/shingi2/0000212500_00236.html [accessed 2024-09-27]

Abbreviations

GIS: geographical information system
KDB: Kokuho database
NDB: National Database of Health Insurance Claims and Specific Health Checkups of Japan

Edited by D Liu; submitted 13.12.24; peer-reviewed by F Zicker, L Jordan, RBL Lim, R Inokuchi; revised version received 16.04.25; accepted 07.06.25; published 06.08.25.

Please cite as:
Morii Y, Nakanishi Y, Nishioka Y, Tsugihashi Y, Noda T, Myojin T, Imamura T, Akahane M
Analyzing Disparity in Geographical Accessibility to Home Medical Care Using a Claims Database and Geographical Information System: Simulation Study
JMIR Aging 2025;8:e70040
URL: https://aging.jmir.org/2025/1/e70040
doi:10.2196/70040

Revisits, Readmission, and Mortality From Emergency Department Admissions for Older Adults With Vague Presentations: Longitudinal Observational Study

Sebastian Alejandro Alvarez Avendano[1], PhD; Amy Cochran[2], PhD; Valerie Odeh Couvertier[1], PhD; Brian Patterson[3], MPH, MD; Manish Shah[3], MPH, MD; Gabriel Zayas-Caban[1], PhD

[1]Department of Industrial and Systems Engineering, College of Engineering, University of Wisconsin-Madison, 3107 Mechanical Engineering Building, 1513 University Avenue, Madison, WI, United States

[2]Population Health Sciences, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI, United States

[3]BerbeeWalsh Department of Emergency Medicine, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Gabriel Zayas-Caban, PhD
Department of Industrial and Systems Engineering, College of Engineering, University of Wisconsin-Madison, 3107 Mechanical Engineering Building, 1513 University Avenue, Madison, WI, United States

Abstract

Background: Older adults (65 years and older) often present to the emergency department (ED) with an unclear need for hospitalization, leading to potentially harmful and costly care. This underscores the importance of measuring the trade-off between admission and discharge for these patients in terms of patient outcomes.

Objective: This study aimed to measure the relationship between disposition decisions and 3-day, 9-day, and 30-day revisits, readmission, and mortality, using causal inference methods that adjust for potential measured and unmeasured confounding.

Methods: A longitudinal observational study (n=3591) was conducted using electronic health records from a large tertiary teaching hospital with an ED between January 1, 2014 and September 27, 2018. The sample consisted of older adult patients with 1 of 6 presentations with significant variability in admission: falls, weakness, syncope, urinary tract infection, pneumonia, and cellulitis. The exposure under consideration was the ED disposition decision (admission to the hospital or discharge). Nine outcome variables were considered: ED revisits, hospital readmission, and mortality within 3, 9, and 30 days of being discharged from either the hospital for admitted patients or the ED for discharged patients.

Results: Admission was estimated to significantly decrease the risk of an ED revisit after discharge (30-day window: −6.4%, 95% CI −7.8 to −5.0), while significantly increasing the risk of hospital readmission (30-day window: 5.8%, 95% CI 5.0 to 6.5) and mortality (30-day window: 1.0%, 95% CI 0.4 to 1.6). Admission was found to be especially adverse for patients with weakness and pneumonia, and relatively less adverse for older adult patients with falls and syncope.

Conclusions: Admission may not be the safe option for older adults with gray area presentations, and while revisits and readmissions are commonly used to evaluate the quality of care in the ED, their divergence suggests that caution should be used when interpreting either in isolation.

(JMIR Aging 2025;8:e55929) doi:10.2196/55929

KEYWORDS

gerontology; geriatric; older adults; elderly; older people; aging; emergency department; emergency room; ED; disposition decision; disposition; discharge; admission; revisit; readmission; observational study; health; hospital

Introduction

Care for acute illnesses has shifted from outpatient offices to emergency departments (EDs), leading to an increase in ED use that has outpaced population growth [1]. EDs diagnose and treat acute illnesses [2]. Therefore, emergency providers diagnose patients, initiate treatment, and predict the disease trajectory to decide whether to admit the patient to an inpatient unit or discharge the patient home [3,4]. Any change to admission decisions can impact outcomes and costs since two-thirds of ED health care costs in the United States come from visits that end in admission [5,6].

Whether to admit or discharge a patient is a weighty decision. This point is best exemplified with older adults (≥65 years of age). Discharging an older adult carries a high risk of adverse health outcomes, especially when compared with younger

patients [7-12]. Hence, increasing interest has been placed on identifying patients at risk for adverse outcomes after ED discharge [7,11,13,14] or on developing strategies for following up on discharged patients [15]. Admitting a patient who can be discharged carries its own risks. Older adults are vulnerable to deconditioning and hospital acquired infection, as well as developing delirium and accelerated cognitive decline [16-19]. These issues underscore the importance of measuring the trade-off between admitting and discharging a patient.

Admission decisions are usually based on well-defined clinical factors and practice guidelines, but often, patients fall into a gray area in which the need for hospital admission is unclear based on objective information or even local standards of care. A large group of such gray-area patients is those presenting with syndromic diagnoses such as falls [20] or weakness [21], or patients presenting with more definite diagnoses that are associated with wide variability in admission decision, such as syncope [22], chest pain [5,23,24], urinary tract infection (UTI) [25], pneumonia [26], and cellulitis [27]. Moreover, other factors influence disposition decisions including triage [28], crowding [29], patient home environment [30], diagnostic testing [31], patient ethnicity [32], and hospital capacity [32]. Consequently, admission rates vary widely for these patients between providers and between hospitals, leading to potentially harmful practice variation [33]. Reducing this variation may avoid harmful admissions while safeguarding patient safety, but in order to do so evidence must be used to guide decision-making.

In this study, we focused on older adults in the ED with gray-area diagnoses as follows: diagnoses associated with clinical ambiguity or high rates of potentially preventable hospitalizations and variability in admissions [16,17]. The goal of this work was to identify how the decision to admit drives subsequent revisit, readmission, and mortality among older adult patients with diagnoses that are syndromic (falls and weakness) or lacking a clear standard of practice (syncope, UTI, pneumonia, or cellulitis). This question was difficult to answer since patients with different disposition decisions differ in their clinical severity, complexity, and needs. Without adjusting for these differences, unfair conclusions can be drawn if patients with different disposition decisions are directly compared. We thus used a causal inference methodology for observational data to measure the relationship between disposition decisions and outcomes for these older adult patients [34,35].

Methods

Data

ED visits were analyzed using electronic health records (EHRs) from a large Midwestern academic health system. The dataset used was from a common EHR system, and the analyzed population consisted of encounters with a specific large ED between January 1, 2014 and September 27, 2018. ED visits were included for older adult patients (65 years of age or older) with presentations for falls (n=1581), weakness (n=564), syncope (n=468), UTI (n=456), pneumonia (n=299), and cellulitis (n=223). For all 6 presentations, inclusion and exclusion criteria were specified to capture patients who, based on objective criteria present in the EHR occupied a “gray area” with regards to criteria for admission. In general, the strategy was to include patients who met diagnostic criteria for each presentation but did not have further abnormalities indicating a clear indication for admission present within discrete fields of the EHR. General inclusion criteria were an acuity level (emergency severity index; ESI) of 2 or 3, excluding visits assigned the most and least severe ESI levels; and treated and either admitted to the hospital or discharged to their residence, excluding patients who eloped, left against medical advice, or were transferred to another facility. General exclusion criteria were a missing disposition decision, a visit in the last 45 days of the sample period (as these patients had inadequate follow-up for outcomes), indicators of acute coronary syndromes (troponin levels >0.10 ng/mL), severe vital sign abnormalities at any point during the ED visit (systolic blood pressure <80 mmHg, respiration rate >30 BPM, pulse oximetry <88%, or heart rate >120 BPM), or a specific diagnosis that unambiguously required admission (ie, stroke, myocardial infarction, or femur fractures). In addition, patients presenting to the ED with a combination of cellulitis and a higher temperature than 100.3 F were eliminated from the analysis since these encounters had a high admission probability.

Variables

Seven variables represented baseline characteristics as follows: age, sex, insurance, history of diabetes, history of congestive heart failure, history of hypertension, and Centers for Medicare and Medicaid Services Hierarchical Condition Category (HCC) score [36]. HCC scores predict health care costs and outcomes by assigning risk weights to comorbidities based on ICD codes; higher scores indicate a greater predicted risk and resource use. These baseline characteristics were selected because they are available prior to the ED visit and can capture key clinical variations that could influence a patient’s ED visit and outcomes.

Five variables represented initial observations serving as a proxy for underlying latent health state at the start of the admission process: acuity, temperature, blood pressure, respiratory rate, and heart rate. Acuity was measured using the ESI level, a 5-level triage system where lower ESI levels (eg, 1 or 2) indicate higher acuity, meaning the patient requires more urgent care and significant resources. These proxy variables were included because they are available before the admission process begins and can be considered indirect (or “noisy”) measurements of the patient’s latent health state.

Two variables represented the admission process: treatment time (duration between when a patient is placed in an ED room after triage for treatment and when their admission decision is made) and admission decision. Patients discharged to a Skilled Nursing Facility or Inpatient Rehabilitation Facility were treated as a discharge.

Three primary outcome variables were considered: ED revisits, hospital readmission, and mortality within 30 days of being discharged from either the hospital for admitted patients or the ED for discharged patients. Binary variables for revisits indicated whether an individual returned to the study ED (based on EHR data) within 30 days of discharge. Similarly, binary variables for readmission and mortality indicated whether an individual was readmitted or died within 30 days of discharge,

respectively. Alternative windows (3 d and 9 d instead of 30 d) were also considered as secondary outcomes, since a short window is likely to capture only subsequent events related to the original ED visit, whereas a long window is likely to capture all subsequent events related to the original ED visit. To optimize this trade-off, a 9-day window has previously been recommended [37].

Missing continuous and categorical variables were imputed with the median and most frequent category, respectively. Continuous variables were standardized and these transformed variables were then used for estimation.

Analysis

Admission decisions were evaluated using 3 types of analyses: unadjusted, adjusted, and subgroup. Each type of analysis was repeated for each of the 3 primary outcomes, and for the overall sample and each presentation group. Unadjusted analyses involved estimating the observed (ie, unadjusted) difference in risk of an outcome between admitted patients and discharged patients. Wald 95% CIs were recovered.

Adjusted analyses, constituting our primary analyses, involved estimating these same risk differences (RDs) but adjusted for the unmeasured or latent health state in addition to measured variables (ie, baseline characteristics and proxy variables). These estimates were recovered by implementing a latent-variable approach for evaluating admission decisions (see Cochran et al [34] for details and applied in Alvarez Avendaño et al [35] to chest pain patients). Briefly, this approach involves modeling both the measured variables and a latent "health state" variable, and then fitting this model to data using expectation maximization, from which an estimate of the average treatment effect (ie, average difference in potential outcomes were we to admit vs discharge a patient) and Wald CIs can be recovered. The latent health state is included to account for potential unmeasured confounding between the admission decision and each outcome. More specifically, it accounts for confounding by indication, whereby admission decisions are based on factors not captured in the data. The inclusion of proxy variables, such as acuity and vitals, is crucial as they provide indirect measures of the latent health state, thereby strengthening the model's ability to mitigate confounding by indication.

The model comprised several regression components: logistic regression to model the latent health state as a function of baseline characteristics; logistic and linear regression to model the proxy variables (eg, acuity and vitals) as a function of latent health and baseline factors; threshold regression to model the admission process (decision and timing) as a function of latent health, baseline characteristics, and proxy variables; and linear regression to model each outcome based on latent health. Once model parameters are estimated, we calculate average treatment effects by comparing outcomes for each visit under hypothetical admission and discharge scenarios while holding all other variables at their original value. We then calculate the difference between these hypothetical outcomes and average this difference across all visits to obtain the overall treatment effect. For subgroup analyses, we recover subgroup-specific average treatment effects by averaging these hypothetical outcomes across only those visits with certain baseline characteristics (eg, female patients).

To check the sensitivity of our conclusions to various factors, several additional analyses were conducted. First, we checked sensitivity to 2 key assumptions for the latent-variable approach, which is that potential outcomes are independent of admission decisions with similar latent health needs and baseline characteristics and that events related to different ED visits are independent. Second, we analyzed the secondary outcomes, which use alternative time windows (3 days and 9 days instead of 30 days). Third, we estimated the same RDs as the main analyses, but only adjusted for the measured variables as opposed to both the measured variables and the latent health state. These estimates were recovered using the causal inference methods known as inverse probability weighting (IPW) and g-estimation.

Due to space considerations, sensitivity analyses are detailed in Multimedia Appendix 1. Importantly, estimates from checking violations to our 2 key assumptions were generally consistent in terms of direction and magnitude with the adjusted estimates reported in the main text, and estimates from IPW and g-estimation were similarly consistent with unadjusted estimates, with a few exceptions detailed in Multimedia Appendix 1. In addition, cellulitis had a small sample size (n=223) relative to other diagnostic groups, leading to imprecise estimates. Therefore, results for cellulitis are presented in Multimedia Appendix 1. Finally, subgroup analyses for individual diagnoses and technical details of all our adjusted analyses can also be found in Multimedia Appendix 1.

Significance was considered at an α level of 0.05. Hypothesis tests were 2-tailed Wald tests. Multiple comparisons were not adjusted for, and as such, nominal CIs and *P* values are reported.

Ethical Considerations

This study was reviewed by the UW Minimal Risk Research Institutional Review Board (ID 2024-0106-CP001) and was deemed to meet the federal criteria for exemption.

Results

Baseline Characteristics

Sample characteristics are summarized in Table 1. Patients were predominantly female (2102/3591, 58.5%) with Medicare insurance (1460/1589, 91.9%) and had an average age of 79 years. Patients were diagnosed with falls (1581/3591, 44%), weakness (564/3591, 16%), syncope (468/3591, 13%), UTI (456/3591, 13%), pneumonia (299/3591, 8%), and cellulitis (223/3591, 6%).

Table . Descriptive statistics of sample by complaint.

	All (n=3591)	Falls (n=1581)	Weakness (n=564)	Syncope (n=468)	UTI[a] (n=456)	Pneumonia (n=299)	Cellulitis (n=223)
Age (years), mean (SD)	79.23 (8.93)	80.56 (8.94)	78.66 (8.4)	76.8 (8.73)	79.22 (8.91)	78.9 (9.27)	76.758 (8.36)
Comorbidity (HCC[b]), mean (SD)	1.65 (1.41)	1.49 (1.22)	1.82 (1.46)	1.32 (1.36)	1.879 (1.6)	2.107 (1.57)	1.965 (1.72)
Heart rate (BPM), mean (SD)	78.07 (14.66)	76.41 (13.54)	78.39 (15.06)	73.11 (13.52)	80.79 (14.82)	87.32 (15.74)	81.41 (14.59)
Temperature (°F), mean (SD)	97.61 (1.4)	97.44 (1.69)	97.56 (0.85)	97.34 (0.72)	97.88 (1.18)	98.55 (1.56)	97.73 (0.8)
Blood pressure (mmHg), mean (SD)	74.2 (13.7)	76.23 (13.92)	74.36 (12.9)	71.91 (12.19)	73.49 (14.72)	69.47 (13.57)	71.99 (12.41)
Respiration rate (BPM), mean (SD)	18.07 (3.18)	17.82 (2.92)	17.86 (3.23)	17.72 (3.57)	18.28 (3.07)	19.8 (3.72)	18.41 (2.58)
Treatment time (hours), mean (SD)	1.82 (0.12)	0.18 (0.1)	0.2 (0.12)	0.22 (0.2)	0.16 (0.07)	0.13 (0.06)	0.15 (0.08)
Female, n (%)	2102 (58.5)	994 (62.9)	299 (53)	255 (54.5)	290 (63.6)	149 (49.8)	115 (51.6)
Insurance, n (%)							
Medicaid/Badger Care	15 (0.4)	3 (0.2)	3 (0.5)	5 (1.1)	2 (0.4)	1 (0.3)	1 (0.4)
Medicare	3278 (91.9)	1460 (92.9)	526 (93.6)	415 (90.8)	409 (90.1)	273 (91.3)	195 (87.4)
Commercial/Worker's Compensation	271 (7.6)	109 (6.9)	33 (5.9)	35 (7.7)	42 (9.3)	25 (8.4)	27 (12.1)
Self-pay	3 (0.1)	0 (0)	0 (0)	2 (0.4)	1 (0.2)	0 (0)	0 (0)
Diabetes, n (%)	744 (20.7)	293 (18.5)	138 (24.5)	81 (17.3)	99 (21.7)	72 (24.1)	61 (27.4)
Congestive Heart failure, n (%)	413 (11.5)	173 (10.9)	62 (11)	33 (7.1)	56 (12.3)	53 (17.1)	36 (16.1)
Hypertension, n (%)	2081 (58.0)	923 (58.4)	350 (62.1)	244 (52.1)	255 (55.9)	170 (56.9)	139 (62.3)
Acuity (=2), n (%)	993 (27.7)	425 (26.9)	139 (24.6)	195 (41.7)	103 (22.6)	106 (35.5)	25 (11.2)
Admitted, n (%)	1401 (39.0)	333 (21.1)	252 (44.7)	179 (38.2)	240 (52.6)	255 (85.3)	142 (63.7)
30-day revisits, n (%)	644 (17.9)	293 (18.5)	108 (19.1)	53 (11.3)	105 (23)	40 (13.4)	45 (20.2)
30-day readmission, n (%)	207 (5.8)	50 (3.2)	36 (6.4)	23 (4.9)	43 (9.4)	33 (11)	22 (9.9)
30-day mortality, n (%)	127 (3.5)	50 (3.2)	30 (5.3)	6 (1.3)	11 (2.4)	27 (9)	3 (1.3)

[a]UTI: urinary tract infection.

[b]HCC: hierarchical condition category.

Revisits

Table 2 summarizes the unadjusted RD between admission and discharge for 30-day revisits across all presentations and for individual presentations. Across all presentations, admission carried a significantly lower unadjusted risk than a discharge of 30-day revisits (RD=−5.2%, 95% CI −6.3 to −4.1). Individual presentations yielded unadjusted estimates that generally agreed with that of the entire sample, with 1 exception. For syncope

patients, admission carried a greater unadjusted risk than a discharge of 30-day revisits (RD=2.5%, 95% CI 0.0 to 5.0).

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RD) for 30-day revisits, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	−5.2 (−6.3 to −4.1)	−6.4 (−7.8 to −5.1)
Falls	−4.5 (−6.5 to −2.5)	−33.2 (−35.5 to −30.8)
Weakness	−8.8 (−11.6 to −6.0)	4.4 (−2.8 to 11.6)
Syncope	2.5 (0 to 5)	−38.6 (−43.2 to −34.1)
UTI[a]	−10.8 (−14.1 to −7.5)	−56.7 (−61.3 to −52.2)
Pneumonia	−3.0 (−7.6 to 1.7)	23.4 (18.7 to 28.1)

[a]UTI: urinary tract infection.

Table 2 also summarizes the estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day revisits. Across all presentations, admission carried a significantly lower adjusted risk than a discharge of 30-day revisits (RD=−6.4%, 95% CI −7.8 to −5.0). Individual presentations yielded adjusted estimates that generally agreed with that of the entire sample, with the following exceptions. For patients with weakness, admission carried a numerically greater adjusted risk than a discharge of 30-day revisits (RD=4.4%, 95% CI −2.8 to 11.6). For patients with pneumonia, admission carried a significantly greater adjusted risk than a discharge of 30-day revisits (RD=23.4%, 95% CI 18.7 to 28.1).

Readmission

Table 3 summarizes the unadjusted RD between admission and discharge for 30-day readmissions across all diagnoses and for individual diagnoses. Across all diagnoses, admission carried a significantly greater unadjusted risk than a discharge of 30-day readmission (RD =14.8%, 95% CI 14.1 to 15.4). Individual diagnoses yielded unadjusted estimates that agreed in terms of direction, significance, and magnitude (within 3%) with that of the entire sample.

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RDs), in percentage points, for 30-day readmissions, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	14.8 (14.1 to 15.4)	5.8 (5.0 to 6.5)
Falls	15.0 (14.1 to 15.9)	3.2 (2.3 to 4.0)
Weakness	14.3 (12.6 to 16.0)	61.6 (57.7 to 65.5)
Syncope	12.8 (11.1 to 14.6)	4.9 (3.0 to 6.9)
UTI[a]	17.9 (15.6 to 20.2)	9.4 (6.8 to 12.1)
Pneumonia	12.9 (8.7 to 17.2)	25.8 (20.9 to 30.6)

[a]UTI: urinary tract infection.

Table 3 also summarizes estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day readmissions. Across all diagnoses, admission carried a significantly greater adjusted risk than a discharge of 30-day readmission (RD=5.8%, 95% CI 5.0 to 6.5). Individual diagnoses yielded adjusted estimates in the same direction as that of the entire sample, but the magnitude was notably larger in a few cases. For patients with weakness, admission carried a greater adjusted risk than a discharge 30-day readmission (RD=61.6%, 95% CI 57.7 to 65.5). For patients with pneumonia, admission carried a significantly greater adjusted risk than a discharge of 30-day readmission (RD=25.8%, 95% CI [20.9, 30.6]).

Mortality

Table 4 summarizes the unadjusted RD between admission and discharge for 30-day mortality across all diagnoses and for individual diagnoses. Across all diagnoses, admission carried a significantly greater unadjusted risk than a discharge of 30-day mortality (RD=3.8%, 95% CI 3.3 to 4.3). Individual diagnoses yielded unadjusted estimates that agreed in terms of direction, significance, and magnitude (within 3%) with those of the entire sample.

Table . Unadjusted and adjusted estimates (95% CI) of risk differences (RDs), in percentage points, for 30-day mortality, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured patient variables.

	Unadjusted RD (95% CI)	Adjusted RD (95% CI)
All	3.8 (3.3 to 4.3)	1.0 (0.4 to 1.6)
Falls	2.8 (1.9 to 3.7)	−0.9 (−1.8 to 0.0)
Weakness	6.9 (5.3 to 8.5)	58.1 (54.1 to 62.0)
Syncope	2.4 (1.6 to 3.3)	−21.9 (−53.6 to 9.7)
UTI[a]	1.9 (0.7 to 3.1)	1.1 (−0.3 to 2.5)
Pneumonia	5.3 (1.4 to 9.2)	7.7 (4.4 to 11.0)

[a]UTI: urinary tract infection.

Table 4 also summarizes estimated RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day mortality. Across all diagnoses, admission carried a significantly greater adjusted risk than a discharge of 30-day mortality (RD=1.0%, 95% CI 0.4 to 1.6). Individual diagnoses yielded adjusted estimates differing from that of the entire sample in several meaningful ways.

Reporting from high to low, admission for patients with weakness carried a significantly greater adjusted risk than a discharge of 30-day mortality (RD=58.1%, 95% CI 54.1 to 62.0). For patients with pneumonia, admission also carried a significantly adjusted risk than a discharge of 30-day mortality (RD=7.7%, 95% CI 4.4 to 11.0). For patients with UTI, admission carried a numerically greater adjusted risk than a discharge of 30-day mortality (RD=1.1%, 95% CI −0.3 to 2.5). For patients with syncope, admission carried a numerically lower adjusted risk of 30-day mortality (RD=−21.9%, 95% CI −53.6 to 9.7). For patients with falls, admission carried a lower adjusted risk than a discharge of 30-day mortality (RD=−0.9%, 95% CI −1.8 to 0.0).

Subgroup Analyses

Table 5 summarizes subgroup-specific estimates of RDs, adjusted for latent health state and measured patient variables, between admission and discharge for 30-day outcomes. All subgroup estimates had the same direction as that of the entire sample: negative for 30-day revisits and positive for 30-day readmission and mortality. For all 30-day outcomes, 5 subgroups had numerically greater adjusted estimates than that of the entire sample. These subgroups (in decreasing order of prominence) were individuals with congenital heart disease, individuals with diabetes, individuals with HCC scores greater than the mean score of 1.65, individuals greater than the mean age of 79.2 years, and individuals with hypertension. Of note, all these subgroups tend to have worse baseline health than their counterpart. The subgroup with the numerically lowest adjusted estimate for all 30-day outcomes was individuals with insurance other than Medicare.

Table . Subgroup-specific adjusted estimates (95% CI) of risk differences, in percentage points, for 30-day outcomes, comparing admission to discharge (reference). Adjusted estimates account for latent health state and measured patient variables.

Variable and subgroup	Revisit	Readmission	Mortality
Age (years)			
65 - 79.2	−7.2 (−8.7 to −5.6)	5.4 (4.7 to 6.2)	0.9 (0.2 to 1.5)
>79.2	−5.6 (−7.0 to −4.1)	6.1 (5.3 to 7.0)	1.2 (0.6 to 1.9)
Sex			
Male	−6.5 (−8.0 to −4.9)	5.8 (5.0 to 6.6)	1.0 (0.4 to 1.7)
Female	−6.4 (−7.9 to −5.0)	5.8 (5.0 to 6.6)	1.0 (0.4 to 1.7)
Insurance			
Medicare	−6.2 (−7.6 to −4.9)	5.8 (5.1 to 6.6)	1.1 (0.5 to 1.7)
Other	−8.5 (−10.8 to −6.3)	4.9 (3.9 to 5.9)	0.6 (−0.2 to 1.3)
Diabetes			
No	−6.8 (−8.2 to −5.4)	5.6 (4.9 to 6.4)	0.9 (0.3 to 1.6)
Yes	−5.0 (−6.7 to −3.2)	6.4 (5.4 to 7.4)	1.4 (0.7 to 2.1)
Congenital heart failure			
No	−7.0 (−8.4 to −5.6)	5.5 (4.8 to 6.2)	0.9 (0.3 to 1.5)
Yes	−1.8 (−3.9 to 0.3)	7.8 (6.6 to 9.0)	2.1 (1.3 to 2.9)
Hypertension			
No	−6.9 (−8.4 to −5.4)	5.6 (4.8 to 6.3)	0.9 (0.3 to 1.5)
Yes	−6.1 (−7.5 to −4.6)	5.9 (5.1 to 6.7)	1.1 (0.5 to 1.7)
Hierarchical condition category			
≤1.65	−7.4 (−8.8 to −6.0)	5.3 (4.6 to 6.1)	0.8 (0.2 to 1.4)
>1.65	−4.7 (−6.2 to −3.2)	6.5 (5.6 to 7.4)	1.4 (0.8 to 2.1)

Discussion

Principal Findings

Our goal was to investigate the impact of admitting older adult patients with 6 common but variably managed ED presentations: falls, weakness, syncope, UTI, pneumonia, and cellulitis. In particular, these conditions show considerable variability in admission practices, and understanding their specific outcomes can help clinicians better determine when admission is beneficial or potentially harmful. Using causal inference methods, we analyzed EHR data to compare outcomes such as ED revisits, hospital readmissions, and mortality, measuring these outcomes at 3, 9, and 30 days after discharge. By focusing on this “gray area” where no standard practice exists, we wanted to provide insights that support more consistent and evidence-based admission decisions for these presentations.

Our primary finding is that admission was associated with a significant decrease in the risk of an ED revisit, but with a significant increase in the risk of hospital readmission and mortality. The association with an increased readmission risk was consistent for all individual diagnoses, for all time frames, and for all estimation approaches (unadjusted, IPW, g-estimation, and latent variable as the primary analysis). The associations with a decreased revisit risk and with an increased mortality risk were generally consistent for individual diagnoses, all time frames, and all approaches, with a few exceptions discussed below. These results would suggest that while admission may help prevent immediate returns to the ED, it could lead to worse long-term outcomes for certain patients. This balance of risks highlights the importance of carefully weighing the short-term benefits of admission against the potential for readmission and mortality, especially in older adults with these conditions.

In our primary analysis, admission was associated with an increased revisit risk for patients with weakness and pneumonia, along with a notably higher readmission and mortality risk for these conditions compared with others. This suggests that patients with weakness or pneumonia may be especially vulnerable to poor outcomes from admission. In contrast, admission was associated with a numerically lower 30-day mortality risk for patients with falls or syncope, indicating that these individuals may benefit more from being admitted. These findings can help guide clinicians in prioritizing which patients may require closer monitoring, alternative care options, or potential admission.

Our subgroup analysis identified patient factors that may influence outcomes from hospital admission. Overall, admission appears less favorable for older individuals (>79 years) and those with comorbidities (ie, congenital heart disease, diabetes, high HCC scores, or hypertension). Important factors for

admission, however, may vary by specific presentation, as shown in our subgroup analysis in Multimedia Appendix 1. Despite these differences, we found consistent trends across subgroups, suggesting that standardized admission recommendations could potentially be based on the type of presentation alone, rather than on patient-specific factors like age, sex, insurance, or comorbidity.

For comparison, another study examined the impact of admitting older adult patients presenting to the ED with chest pain, another presentation with variable admission practices, and found that admission had a lower readmission risk but an increased mortality risk [35]. This study similarly estimated an increased mortality risk, yet it also found an increase in readmission risk for almost all presentations. The association of admission with both increased and decreased risks for revisits, readmission, and mortality, underscores why standardized care pathways are lacking for these presentations [21]. While our results suggest that, in general, discharge may be associated with lower readmission and mortality risks, discharging an older adult can still carry significant risks. These findings support the need for safe alternatives to hospitalization, such as outpatient care coordination or at-home ambulatory providers [38]. Such alternatives have been shown to effectively reduce hospital readmission rates on older adult patients discharged from the ED [39,40].

Limitations

A primary limitation of this study is the potential for unmeasured confounding, especially confounding by indication, where treatment decisions are influenced by clinical indicators not fully captured in the EHR. This limitation is significant because it directly impacts our ability to interpret the findings causally. Although we used multiple methods, including a latent variable approach to account for unobserved health status [34], no method can fully guarantee that all confounding is addressed. Consequently, our findings remain subject to potential bias if residual confounding exists. This is particularly relevant for interpreting the effect of admission on outcomes, as patients who are “sicker” based on unmeasured factors are more likely to be admitted, and these same factors could drive poor outcomes. While the latent variable approach yielded estimates generally more favorable to admission—suggesting some adjustment for underlying severity—and results were consistent across different time frames and presentations, caution is warranted in interpreting these effects as causal. This limitation highlights the need for further research to strengthen the causal inference of admission effects on outcomes.

In addition, this was a retrospective, observational study conducted using EHR data from an academic ED in the US Midwest, thereby limiting the generalizability of our findings. Moreover, the resulting patient population for specific presentations after applying our filters was relatively small, leading to nonsignificant estimates. Replicating this analysis in different EDs with larger samples is needed.

Conclusions

This study is one of the few to quantify the risks and benefits associated with admission decisions for ED presentations associated with unclear need for hospitalization. Existing studies have mainly focused on identifying factors that influence hospital admission of noncritically ill patients, often through physician surveys [41] or identifying risk patterns [33,42,43]. This study is intended to quantify the trade-offs of admission at a population level rather than directly influencing bedside decisions. Our findings indicate that admission is not always the “safe” choice, underscoring the need for further research to establish specific criteria that can better guide admission decisions for older adults in the ED.

Acknowledgments

This research was supported by the Clinical and Translational Science Award (CTSA) program, through the NIH National Center for Advancing Translational Sciences (NCATS) (GZC, grant numbers UL1TR002373 and KL2TR002374); and by the University of Wisconsin-Madison School of Medicine and Public Health's Wisconsin Partnership Program (WPP) (VOC, grant number 4358). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or WPP. The authors also wish to thank the Health Innovation Program at the University of Wisconsin-Madison School of Medicine and Public Health for their technical expertise and support in preparing the final curated dataset.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Accompanying appendix reporting (1) IPW and g-estimation results, (2) description of the latent variable approach, (3) 3-day outcomes results, (4) sensitivity analyses, (5) subgroup analyses for individual diagnoses, (6) analysis for cellulitis diagnosis, and (7) analysis for syndromic and no clear standard practice diagnoses.
[DOCX File, 430 KB - aging_v8i1e55929_app1.docx]

References

1. Pitts SR, Carrier ER, Rich EC, Kellermann AL. Where Americans get acute care: increasingly, it’s not at their doctor’s office. Health Aff (Millwood) 2010 Sep;29(9):1620-1629. [doi: 10.1377/hlthaff.2009.1026]

2. Kellermann AL, Martinez R. The ER, 50 years on. N Engl J Med 2011 Jun 16;364(24):2278-2279. [doi: 10.1056/NEJMp1101544]
3. Tang N, Stein J, Hsia RY, Maselli JH, Gonzales R. Trends and characteristics of US emergency department visits, 1997-2007. JAMA 2010 Aug 11;304(6):664-670. [doi: 10.1001/jama.2010.1112] [Medline: 20699458]
4. Morganti KG, Bauhoff S, Blanchard JC, et al. The Evolving Role of Emergency Departments in the United States: Rand health quarterly The RAND Corporation; 2013, Vol. 3.
5. Sabbatini AK, Nallamothu BK, Kocher KE. Reducing variation in hospital admissions from the emergency department for low-mortality conditions may produce savings. Health Aff (Millwood) 2014 Sep;33(9):1655-1663. [doi: 10.1377/hlthaff.2013.1318]
6. Galarraga JE, Pines JM. Costs of ED episodes of care in the United States. Am J Emerg Med 2016 Mar;34(3):357-365. [doi: 10.1016/j.ajem.2015.06.001]
7. Aminzadeh F, Dalziel WB. Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions. Ann Emerg Med 2002 Mar;39(3):238-247. [doi: 10.1067/mem.2002.121523]
8. Lowenstein SR, Crescenzi CA, Kern DC, Steel K. Care of the elderly in the emergency department. Ann Emerg Med 1986 May;15(5):528-535. [doi: 10.1016/S0196-0644(86)80987-8]
9. Hedges JR, Singal BM, Rousseau EW, et al. Geriatric patient emergency visits part II: perceptions of visits by geriatric and younger patients. Ann Emerg Med 1992 Jul;21(7):808-813. [doi: 10.1016/S0196-0644(05)81026-1]
10. Denman SJ, Ettinger WH, Zarkin BA, Coon PJ, Casani JA. Short‐term outcomes of elderly patients discharged from an emergency department. J Am Geriatr Soc 1989 Oct;37(10):937-943. [doi: 10.1111/j.1532-5415.1989.tb07278.x]
11. Caplan GA, Brown A, Croker WD, Doolan J. Risk of admission within 4 weeks of discharge of elderly patients from the emergency department—the DEED study. Age Ageing 1998;27(6):697-702. [doi: 10.1093/ageing/27.6.697]
12. Rowland K, Maitra AK, Richardson DA, Hudson K, Woodhouse KW. The discharge of elderly patients from an accident and emergency department: functional changes and risk of readmission. Age Ageing 1990;19(6):415-418. [doi: 10.1093/ageing/19.6.415]
13. Martin-Gill C, Reiser RC. Risk factors for 72-hour admission to the ED. Am J Emerg Med 2004 Oct;22(6):448-453. [doi: 10.1016/j.ajem.2004.07.023]
14. Gerson LW, Rousseau EW, Hogan TM, Bernstein E, Kalbfleisch N. Multicenter study of case finding in elderly emergency department patients. Acad Emerg Med 1995 Aug;2(8):729-734. [doi: 10.1111/j.1553-2712.1995.tb03626.x]
15. Stuck AR, Crowley C, Killeen J, Castillo EM. National survey of emergency physicians concerning home-based care options as alternatives to emergency department–based hospital admissions. J Emerg Med 2017 Nov;53(5):623-628. [doi: 10.1016/j.jemermed.2017.05.036]
16. Kozak LJ, Hall MJ, Owings MF. Trends in avoidable hospitalizations, 1980–1998. Health Aff (Millwood) 2001 Mar;20(2):225-232. [doi: 10.1377/hlthaff.20.2.225]
17. Nyweide DJ, Anthony DL, Bynum JPW, et al. Continuity of care and the risk of preventable hospitalization in older adults. JAMA Intern Med 2013 Nov 11;173(20):1879-1885. [doi: 10.1001/jamainternmed.2013.10059] [Medline: 24043127]
18. Hussain M, Oppenheim BA, O'Neill P, Trembath C, Morris J, Horan MA. Prospective survey of the incidence, risk factors and outcome of hospital-acquired infections in the elderly. J Hosp Infect 1996 Feb;32(2):117-126. [doi: 10.1016/S0195-6701(96)90053-5]
19. Taylor ME, Oppenheim BA. Hospital-acquired infection in elderly patients. J Hosp Infect 1998 Apr;38(4):245-260. [doi: 10.1016/S0195-6701(98)90074-3]
20. Hartholt KA, van der Velde N, Looman CWN, et al. Trends in fall-related hospital admissions in older persons in the Netherlands. Arch Intern Med 2010 May 24;170(10):905-911. [doi: 10.1001/archinternmed.2010.106] [Medline: 20498419]
21. Eagles D, Stiell IG, Clement CM, et al. International survey of emergency physicians' priorities for clinical decision rules. Acad Emerg Med 2008 Feb;15(2):177-182. [doi: 10.1111/j.1553-2712.2008.00035.x]
22. Daccarett M, Jetter TL, Wasmund SL, Brignole M, Hamdan MH. Syncope in the emergency department: comparison of standardized admission criteria with clinical practice. Europace 2011 Nov 1;13(11):1632-1638. [doi: 10.1093/europace/eur201]
23. Kocher KE, Meurer WJ, Fazel R, Scott PA, Krumholz HM, Nallamothu BK. National trends in use of computed tomography in the emergency department. Ann Emerg Med 2011 Nov;58(5):452-462. [doi: 10.1016/j.annemergmed.2011.05.020]
24. Kocher KE. Achieving the holy grail of emergency department evaluation for chest pain. Circ Cardiovasc Qual Outcomes 2017 Oct;10(10). [doi: 10.1161/CIRCOUTCOMES.117.004026]
25. Sammon JD, Sharma P, Rahbar H, et al. Predictors of admission in patients presenting to the emergency department with urinary tract infection. World J Urol 2014 Jun;32(3):813-819. [doi: 10.1007/s00345-013-1167-3]
26. Venkatesh AK, Chou SC, Li SX, et al. Association between insurance status and access to hospital care in emergency department disposition. JAMA Intern Med 2019 May 1;179(5):686-693. [doi: 10.1001/jamainternmed.2019.0037] [Medline: 30933243]
27. Volz KA, Canham L, Kaplan E, Sanchez LD, Shapiro NI, Grossman SA. Identifying patients with cellulitis who are likely to require inpatient admission after a stay in an ED observation unit. Am J Emerg Med 2013 Feb;31(2):360-364. [doi: 10.1016/j.ajem.2012.09.005]

28. Calder LA, Forster AJ, Stiell IG, et al. Mapping out the emergency department disposition decision for high-acuity patients. Ann Emerg Med 2012 Nov;60(5):567-576. [doi: 10.1016/j.annemergmed.2012.04.013]
29. Gorski JK, Batt RJ, Otles E, Shah MN, Hamedani AG, Patterson BW. The impact of emergency department census on the decision to admit. Acad Emerg Med 2017 Jan;24(1):13-21. [doi: 10.1111/acem.13103] [Medline: 27641060]
30. Nugus P, Carroll K, Hewett DG, Short A, Forero R, Braithwaite J. Integrated care in the emergency department: a complex adaptive systems perspective. Soc Sci Med 2010 Dec;71(11):1997-2004. [doi: 10.1016/j.socscimed.2010.08.013]
31. Florin TA, French B, Zorc JJ, Alpern ER, Shah SS. Variation in emergency department diagnostic testing and disposition outcomes in pneumonia. Pediatrics 2013 Aug;132(2):237-244. [doi: 10.1542/peds.2013-0179] [Medline: 23878049]
32. Studnicki J, Platonova EA, Fisher JW. Hospital-level variation in the percentage of admissions originating in the emergency department. Am J Emerg Med 2012 Oct;30(8):1441-1446. [doi: 10.1016/j.ajem.2011.11.008]
33. Pines JM, Mutter RL, Zocchi MS. Variation in emergency department admission rates across the United States. Med Care Res Rev 2013 Apr;70(2):218-231. [doi: 10.1177/1077558712470565]
34. Cochran AL, Rathouz PJ, Kocher KE, Zayas‐Cabán G. A latent variable approach to potential outcomes for emergency department admission decisions. Stat Med 2019 Sep 10;38(20):3911-3935. [doi: 10.1002/sim.8210]
35. Alvarez Avendaño SA, Cochran AL, Patterson B, Shah M, Smith M, Zayas-Cabán G. The average effect of emergency department admission on readmission and mortality for older adults with chest pain. Med Care 2020;58(10):881-888. [doi: 10.1097/MLR.0000000000001375]
36. Pope GC, Ellis RP, Ash AS, et al. Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment: Health Economics Research; 2000.
37. Rising KL, Victor TW, Hollander JE, Carr BG. Patient returns to the emergency department: the time-to-return curve. Acad Emerg Med 2014 Aug;21(8):864-871. [doi: 10.1111/acem.12442]
38. Kilaru AS, Resnick D, Flynn D, et al. Practical Alternative to Hospitalization for Emergency Department Patients (PATH): A Feasibility Study: Healthcare Elsevier; 2021:100545. [doi: 10.1016/j.hjdsi.2021.100545]
39. Guttman A, Afilalo M, Guttman R, et al. An emergency department–based nurse discharge coordinator for elder patients: does it make a difference? Acad Emerg Med 2004 Dec;11(12):1318-1327. [doi: 10.1197/j.aem.2004.07.006]
40. Caplan GA, Williams AJ, Daly B, Abraham K. A randomized, controlled trial of comprehensive geriatric assessment and multidisciplinary intervention after discharge of elderly from the emergency department—the DEED II study. J Am Geriatr Soc 2004 Sep;52(9):1417-1423. [doi: 10.1111/j.1532-5415.2004.52401.x]
41. Lewis Hunter AE, Spatz ES, Bernstein SL, Rosenthal MS. Factors influencing hospital admission of non-critically Ill patients presenting to the emergency department: a cross-sectional study. J Gen Intern Med 2016 Jan;31(1):37-44. [doi: 10.1007/s11606-015-3438-8] [Medline: 26084975]
42. Oster A, Bindman AB. Emergency department visits for ambulatory care sensitive conditions: insights into preventable hospitalizations. Med Care 2003;41(1):198-207. [doi: 10.1097/01.MLR.0000045021.70297.9F]
43. Pappas G, Hadden WC, Kozak LJ, Fisher GF. Potentially avoidable hospitalizations: inequalities in rates between US socioeconomic groups. Am J Public Health 1997 May;87(5):811-816. [doi: 10.2105/AJPH.87.5.811]

Abbreviations

ED: emergency department
EHR: electronic health record
ESI: emergency severity index
HCC: hierarchical condition category
IPW: inverse probability weighting
RD: risk difference
UTI: urinary tract infection

Edited by Y Jiang; submitted 29.12.23; peer-reviewed by A Hassan, K Musikatavorn, V Winter; revised version received 04.11.24; accepted 05.12.24; published 06.02.25.

Please cite as:
Alvarez Avendano SA, Cochran A, Odeh Couvertier V, Patterson B, Shah M, Zayas-Caban G
Revisits, Readmission, and Mortality From Emergency Department Admissions for Older Adults With Vague Presentations: Longitudinal Observational Study
JMIR Aging 2025;8:e55929
URL: https://aging.jmir.org/2025/1/e55929
doi:10.2196/55929

Original Paper

Passive Remote Monitoring Technologies' Influence on Home Care Clients' Ability to Stay Home: Multiprovincial Randomized Controlled Trial

Lorie Donelle[1], PhD; Bradley Hiebert[2,3], PhD; Grace Warner[4], PhD; Michael Reid[5], MSc, MMM; Jennifer Reid[6], MSc; Salimah Shariff[2,6], PhD; Emily Richard[7], PhD; Sandra Regan[2], PhD; Lori Weeks[8], PhD; Kathleen Ledoux[2], PhD

[1]College of Nursing, University of South Carolina, Columbia, SC, United States
[2]Arthur Labatt Family School of Nursing, Western University, London, ON, Canada
[3]St Joseph's Health Care London, London, ON, Canada
[4]School of Occupational Therapy, Dalhousie University, Halifax, NS, Canada
[5]Department of Health and Epidemiology, Dalhousie University, Halifax, NS, Canada
[6]ICES Western, London Health Sciences Centre Research Institute, London, ON, Canada
[7]Faculty of Nursing, University of New Brunswick, Moncton, NB, Canada
[8]School of Health Administration, Dalhousie University, Halifax, NS, Canada

Corresponding Author:
Lorie Donelle, PhD
College of Nursing
University of South Carolina
1601 Greene St
Columbia, SC, 29208
United States
Phone: 1 803 777 6528
Email: ldonelle@mailbox.sc.edu

Abstract

Background: Researchers in Nova Scotia and Ontario, Canada, implemented a passive remote monitoring (PRM) model of home care unique to their health system contexts. Each PRM model integrated tailored PRM devices (eg, motion sensors, cameras, and door alarms) into home care patients' residences with the aim of linking patients, family and friend caregivers, and health care providers to support older adults' aging in place.

Objective: The purpose of this study was to examine the use of PRM technologies in the home to support older adults' safe aging in place and avoidance or delay of higher levels of care.

Methods: This multiprovincial pragmatic randomized controlled trial examined how PRM technologies support older adults to safely remain in their home and avoid or delay admission to higher levels of care. Pairs of home care patients and their family and friend caregivers were recruited in Ontario and Nova Scotia. Participant pairs were randomly assigned to one of two conditions: (1) standard home care (ie, control) or (2) standard home care plus study-provided PRM (ie, intervention). Participants provided their provincial health insurance numbers to link with provincial health administrative databases and identify if patients were admitted to higher levels of care after 1 year. Cox proportional hazards models were used to evaluate the primary outcome in each province.

Results: In total, 313 patient-caregiver pairs were recruited: 174 pairs in Ontario (intervention: n=60; control: n=114) and 139 pairs in Nova Scotia (intervention: n=45; control: n=94). Results indicate PRM was associated with a nonsignificant 30% reduction in risk of patients being admitted to higher levels of care in Ontario (hazard ratio 0.7, 95% CI 0.3-1.4) and no reduction in risk in Nova Scotia (hazard ratio 1.1, 95% CI 0.3-3.7). Adjusting for patient sex had no impact on model estimates for either province.

Conclusions: Limitations related, in part, to the impact of the COVID-19 pandemic may have contributed to the effectiveness of the intervention. While our study did not yield statistically significant results (P=.30 and P=.90) regarding the effectiveness of the PRM model in prolonging home stays, the observed trends suggest that technology-assisted aging in place may be a valuable

goal for older adults. Further study is required to understand if longer follow-up time allows more effects of PRM on patients' avoidance of higher levels of care to be detected.

Trial Registration: ISRCTN ISRCTN79884651; https://www.isrctn.com/ISRCTN79884651

International Registered Report Identifier (IRRID): RR2-10.2196/15027

(JMIR Aging 2025;8:e69107) doi:10.2196/69107

KEYWORDS

remote monitoring technology; home care; health service use; aging in place

Introduction

Creative and effective ways to meet the health care needs of an aging population, often related to the high prevalence of chronic disease, are complicated by a strained health care work force, competing economic priorities, and evolving health information. Yet, these challenges create opportunities to reimagine how health care systems can provide Canadians with the right care, at the right time, in the right place. Although older adults may experience significant chronic disease that may challenge their ability to live independently, most older adults want to remain at home [1,2]. This sentiment was underscored by the findings of a recent survey where 85% of Canadian adults and 96% of Canadians aged ≥65 years reported that they would do everything they can to avoid institutionalization in a long-term care (LTC) setting [2,3]. Home care services that are appropriately managed and integrated into the health care system support older adults to successfully age at home, can improve the health and well-being of older adults and their families, and reduce costs of care in hospitals and LTC facilities [2,4]. However, in Canada, beyond the challenges previously noted, the provision of publicly funded home care services is limited and home care providers are increasingly unable to meet existing demands for home care service needs [2,4-6]. In fact, approximately 75% of home care support for older Canadians is provided *supposedly for free* by an unpaid family and friend caregiver [2,7], creating enhanced personal risk of physical health problems, stress, burnout, and depression [2,8].

A noted challenge of caring for older adults at home is providing suitable home care services and ensuring that older adults can safely follow established treatment plans. Common adverse events among older adults at home (eg, medication administration errors and falls) have been associated with increased use of health services, disability, and death [9,10]. Furthermore, many older adults have complex health needs, and the home care service gaps for older adults with complex care needs are further magnified by limited health human resources, access to partial home care services, a lack of direct support for or involvement of caregivers, and a lack of innovative strategies to expand home care [2,5]. These challenges highlight the urgency to develop more effective and respectful home care options for older adults.

A plausible intervention to address these home care challenges is the implementation of active and passive remote monitoring (PRM) technologies to support older adults to remain in their homes [11]. Active monitoring applications require individual participation, such as pushing a button on a wearable device (eg, pendant or bracelet), while passive monitoring technologies, such as motion sensors, do not require any action by the individual for the system to work. Remote monitoring is useful for tracking the behaviors of older adults with cognitive decline (eg, forgetting to take medications) and intervening quickly in the case of adverse events, such as falls. It also increases individuals' confidence in their ability to care for themselves and live independently at home [11,12]. These technologies can also benefit older adults living at home and their caregivers by increasing communication and collaboration among people in their circle of care and by enabling big data analytics that can contribute to improving health care delivery practices [11,12]. While there is emerging evidence around the value of monitoring technologies to support older adults at home, there is limited insight into the outcomes of PRM, the preferences of older adults and their caregivers, and the costs and benefits associated with the wide variety of remote monitoring technologies available.

The purpose of this study was to examine the use of PRM technologies in the home to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. We evaluated a technology-enabled remote monitoring model of home care conceptually informed by the principles of person-centered care, family as a client, supported self-management, and stakeholder collaboration, to address gaps in home care for older adults. Supportive self-management speaks to the extended health care responsibilities taken up by older adults and their caregivers that create care work. Supportive self-management within home care would include both clinical (eg, medication administration, symptom monitoring, and care coordination) and nonclinical (eg, homemaking, meal preparation, supportive housing, daily check-ins, transportation, and a 24/7 helpline) funded services [13,14]. In this model of care, family was conceptualized as the patient or client receiving home care services *and* their caregiver. Typically, caregivers play an integral role as part of the care team to support older adults in the home; health care assessments remain largely focused on the needs of the older adult with relatively marginal inquiry into caregiver health and well-being [13,14]. The primary focus of this study was to investigate the impact of the PRM model of home care on older adult home care recipients with a secondary focus on assessing the impact of PRM on caregivers' health and well-being. Reported here are our findings as to whether PRM along with usual home care supported older adults with complex care needs to safely remain in their home longer and delay or avoid admission to higher levels of care (ie, hospitalization and LTC)

when compared with older adults receiving usual home care alone.

Methods

Overview

This study was informed by the findings of previous pilot studies conducted in the Canadian provinces of Alberta, British Columbia, New Brunswick, and Nova Scotia [15,16]. The primary outcome of this study was to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. This outcome was operationalized among complex care older adult home care recipients as the occurrence of and time to the following events: terminal admission to the hospital awaiting admission to LTC and direct admission to LTC. The secondary outcomes were an assessment of health service use and mortality rates of home care recipients within 1 year of trial enrollment. A complete description of the study's methodology is reported elsewhere [17] and is summarized in the subsequent sections.

Study Design

This study was an unblinded pragmatic randomized controlled trial (PRCT) [18,19] with 2 parallel arms in two Canadian provincial sites: (1) Central, Western, and Northern Zones in Nova Scotia and (2) the South West Region in Ontario. PRCTs are well-suited to supporting decision-making around complex interventions tested in the *real world* in comparison to usual forms of care [20,21]. As PRCTs are meant to capture real-world situations and the experiences of individuals, broad inclusion criteria developed in collaboration with regional health care partners were purposely selected for this study (see Textbox 1 for the full criteria). For this study, participants were recruited in home care client-caregiver dyads because the PRM requires that older adults have a family caregiver who is willing to receive system notifications. This PRCT is registered with the ISRCTN (registration 79884651).

Textbox 1. Participant eligibility.

Participant group and inclusion criteria (eligible if all criteria are met)

- Home care clients
- Adult who is aged ≥65 years
- Requires home care and is at risk for higher levels of care as determined by the home care provider who makes these decisions
- Has a caregiver who is willing and able to receive the remote monitoring sensor notifications using a cell phone or a regular phone (ie, landline)
- Able to read and write in English or French
- Has the decisional capacity to consent or have a substitute decision-maker to consent for participation
- Caregivers
- Adult who is aged ≥18 years
- Is a caregiver to an adult who is aged ≥65 years that requires home care services and is assessed by home care service providers to be at risk for higher levels of care
- Can be a spouse, partner, child, sibling, other family relation, or friend who helps care for the patient at home
- Willing and able to receive the remote monitoring sensor notifications using a cell phone or a regular phone (ie, landline)
- Able to read and write in either English or French
- Has the decisional capacity to consent for participation

Recruitment

The participant sample size was determined using information from one health region regarding the time to higher levels of care for older adults who previously required complex home care. The sample size calculation was based on the following criteria: total institutionalization proportion among controls being 0.41, and the proportion for the experimental participants being 0.27 (a 34% reduction compared with controls); a 10% attrition rate due to dropout or loss to follow-up; a power of 80%; and a statistical significance level of α=.05. This resulted in a total target sample size of 160 home care clients (and paired caregivers) for the intervention group and 320 home care clients (and paired caregivers) in the control group for a total study sample size of 480 home care clients and 480 caregivers across the Ontario and Nova Scotia study sites. The required sample was estimated based on a prospective test of independence (continuity-corrected chi-square statistic test) between the experimental and control groups using 2 controls per intervention case to increase the study's power due to the cost of the intervention [22].

Participant recruitment occurred from April 2017 through January 2020 and was supported by the respective provincial regional authorities. Assessors and care coordinators employed at the health service provider that assesses and facilitates home care services in each region identified potential study participants (ie, clients) based on the study inclusion criteria (Textbox 1). The home care clients may have been on the assessors and care coordinators' case load for up to 6 months before being contacted about the study. Care coordinators provided information about the study to potential study participants and their family and friend caregivers and requested permission to provide their names and contact information to

the study research coordinator. Research staff contacted potential participants who agreed to have their contact information disclosed to arrange an appointment (in person or by telephone) to provide more information about the study. If both the home care client and family and friend caregiver met the eligibility criteria and verbally agreed to participate, the research staff scheduled an in-home meeting with the pair to obtain written consent and complete the baseline data collection forms (detailed in the *Data Collection* section).

Randomization

With participants' consent, baseline data were collected before random allocation into control or intervention groups. A block randomization process was used after every 3 pairs of participants were recruited [23] to randomly assign pairs in a 2:1 ratio to usual home care (ie, control group) or to usual home care plus study-provided PRM (ie, intervention group). This process was carried out independently in each province by members of the research team. Allocation bias was addressed through allocation concealment, and neither the home care case coordinator nor the researchers knew which group the participant would be assigned before baseline data collection. Following randomization research staff contacted each participant pair to inform them of their study arm. For intervention participants, research staff also forwarded their contact information to the technology provider to schedule an installation date.

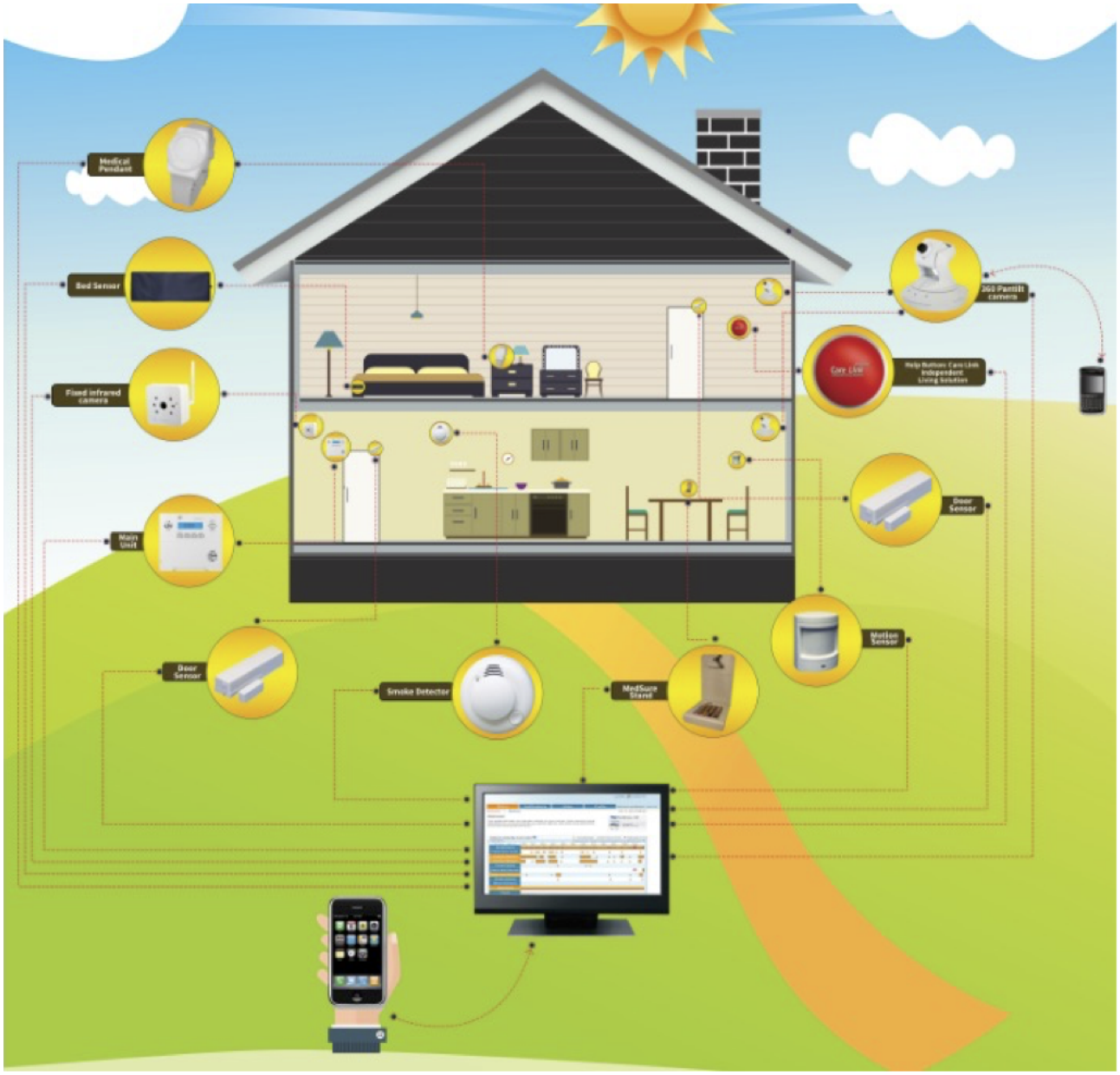
Control Group

Participants received their usual publicly funded home care services provided by their provincial or regional health care authority; some participants also purchased privately funded home care services to supplement the publicly funded services received. These services included home visits by assistive personnel for activities of daily living, nursing care, and other supports deemed necessary by the home care case coordinator or case manager. Home care assessments were conducted as usual by the regional authority case coordinator or case manager. Once the assessment was completed and care services were decided, services were provided by a contracted home care agency.

Intervention Group

Participants received their usual home care services (described earlier) and PRM systems provided and funded by the study technology partner. Once enrolled in the study, the technology partner visited the clients and their caregiver. Participants received written and verbal overviews of the PRM options. Each PRM system installed was tailored to meet the needs and preferences of both the home care client and their caregiver. PRM system options included a combination of long-life battery–powered sensors including motion sensors; Wi-Fi–enabled cameras; fall alert pendants or bracelets; contact sensors for internal or external doors, cupboards, and refrigerators; pressure mats that could be placed at the base of a toilet, under a mattress, or on a chair; and a medication sensor system. Each PRM system had a main panel that was connected to a phone jack installed in the home care client's residence. The main panel received the signals from each sensor and sent them to a secured server to be transmitted back to the authorized recipient (eg, a caregiver) via a cellular signal or broadband module (Global System for Mobile Communications radio). The PRM systems only required internet access to transmit Wi-Fi-camera data. See Figure 1 for an example of how a PRM system used in this study could be set up.

Figure 1. Example of a passive remote monitoring system setup.



The role of the caregiver was key to the successful implementation of the PRM intervention. In addition to an array of sensors, the technology partner, in consultation with the home care client and caregiver, configured each PRM system to provide home care clients and their caregivers with a range of alerts for atypical events tailored to their situation. Examples of alerts include reminders for clients to take their medication; the ability for caregivers to monitor medication use; home care client movement patterns within the home or when leaving the home; fall emergencies; abnormal eating patterns; and abnormal length of time in bed, on a chair, or in a toilet. Notifications of atypical events were sent to the caregiver via email, SMS text message, or phone call. Notably, notifications of home care clients' atypical behaviors were not directly transmitted to a health care provider; rather, health care providers were notified at the discretion of the caregiver and in collaboration with the home care client, as appropriate. Possible caregiver actions based on notifications could include a telephone call to check on the client's safety, deploying assistive home care supports, checking the video (if cameras are present), taking the client to seek medical attention, or emergency action such as calling an ambulance. The technology partner provided education (written and verbal) to the home care clients and caregivers about PRM sensors and the types of alerts; the technology provider technician communicated any changes to the sensor system within the home with the research staff.

The intervention was provided to participants for 12 months during the study. After this time, the home care client was transitioned to usual care if they did not wish to retain their PRM system at a reduced cost.

Data Collection

Data Collection Window

Data collection aligned with the rolling recruitment of participants (see the *Recruitment* section) and lasted from April 2017 to February 2021. Paper format surveys comprised of standardized and researcher-developed tools were used to collect

data from participant pairs three times over the course of a 12-month period: (1) at the point of recruitment when consent was obtained (ie, baseline), (2) 6 months following recruitment, and (3) 12 months following recruitment. Data were collected in person at the home care client's home at baseline and by telephone or in person at the client's home for the 6-month and 12-month follow-ups. Home care clients and caregivers were interviewed separately where possible.

Survey

Both home care clients and caregivers completed a survey consisting of validated questionnaires and researcher-developed questions to provide demographic information (eg, date of birth, education, sex, marital status, and household income); provincial health insurance numbers (used for data linkage to provincial administrative health databases); feelings of home care clients' safety in their home; appraisals of home care quality; and satisfaction with and impact of PRM on both client and caregiver well-being (asked to the intervention group only). Additional researcher-developed items were asked of caregivers to understand their life satisfaction, self-rated mental health, and daily stress.

Home care clients completed standardized tools to assess their quality of life [24], Hospital Admission Risk Profile score, cognitive capacity (Mini-Mental State Exam) [25], and independent activities of daily living [26]. Caregivers also responded to standardized tools to assess their caregiving context, experiences, and information needs [27,28]; caregiver burden [29]; the home care client's abilities to complete independent activities of daily living (a part of the Hospital Admission Risk Profile) [26]; positive aspects of caregiving [30]; presenteeism at work [31]; and the financial impact of caregiving [32].

Impact of the COVID-19 Pandemic on Data Collection

The onset of the COVID-19 pandemic response in March 2020 occurred during the data collection period. Research staff adhered to all public health guidelines for risk mitigation and collected data entirely by phone during this period. Changes to health care service delivery models and stay-at-home orders in Ontario and Nova Scotia altered how these participants (1) could access and use health care services and (2) interacted with their participant partner (ie, the other half of the client-caregiver dyad). Participants' adherence to public health stay-at-home orders resulted in behaviors (eg, hesitancy to engage with acute and LTC services) that completely conflicted with and was contrary to our research outcome creating significant bias within the proportion of participants' data collected after March 2020. In consultation with epidemiologists and statisticians, a decision was made to control for the effects of the pandemic response, and participant pairs were only included for analysis if all study activities were completed before March 2020.

Administrative Data Linkage

Overview

Participants were enrolled between April 11, 2017, and January 15, 2020, in Ontario and between November 29, 2017, and December 16, 2019, in Nova Scotia. To determine the impact of the intervention on the primary study outcomes, the data collected by the research team (herein referred to as the PRCT database) was linked to health administrative databases held at the Institute for Clinical Evaluative Sciences (ICES; Ontario) and Health Data Nova Scotia (HDNS; Nova Scotia).

ICES is an independent, nonprofit research institute funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. As a prescribed entity under Ontario's privacy legislation, ICES is authorized to collect and use health care data for the purposes of health system analysis, evaluation, and decision support. Secure access to these data is governed by policies and procedures that are approved by the Information and Privacy Commissioner of Ontario.

HDNS is a similar institution to ICES. It is a data repository based in the Faculty of Medicine, Department of Community Health and Epidemiology at Dalhousie University, that is focused on supporting data-driven research for researchers in Nova Scotia. HDNS facilitates research and innovation in Nova Scotia by providing access to linkable administrative health data and analysis for research and health service assessment purposes in a secure, controlled environment, while respecting the privacy and confidentiality of Nova Scotians.

Data Sources

We used the PRCT database to identify patient and caregiver participants, including measures of risk that the patient had to alternate levels of care and hospital admission and measures of burden on caregivers. Health administrative databases in both Ontario and Nova Scotia were used to determine other baseline characteristics and outcomes and identify participant data, as follows:

1. Demographic and geographic information as recorded in administrative data
2. Preexisting comorbidities (using validated algorithms)
3. Previous emergency department visits
4. Previous surgeries
5. Previous outpatient, primary care, and specialist visits and follow-up LTC claims
6. Previous prescription use and follow-up LTC
7. Previous home care services
8. Previous and follow-up hospital admissions
9. International classification of diseases diagnosis codes—version 10 and Ontario health insurance plan diagnosis codes.

In Ontario, datasets were linked using unique encoded identifiers and analyzed at ICES. In Nova Scotia, the databases were matched whenever possible. When databases could not be matched, HDNS used alternate databases with similar data to approximate what was done in Ontario. For a complete listing of all databases used, see Table 1.

Table 1. Databases used by the Institute for Clinical Evaluative Sciences (ICES) and Health Data Nova Scotia (HDNS) to support analysis.

Characteristics	ICES databases (Ontario)	HDNS databases (Nova Scotia)
Demographic and geographic information	The registered persons database and PCCF[a]	• Insured patient registry (MASTER) • Postal code conversion file (PCCF+)
Preexisting comorbidities using validated algorithms	Ontario Diabetes Dataset, Chronic Obstructive Pulmonary Disease, congestive heart failure, Ontario hypertension dataset (HYPER), Ontario asthma dataset (ASTHMA), Ontario dementia database (DEMENTIA)	• MSI physician's billings (MED)
Emergency department visits	NACRS[b]	• NACRS
Surgeries	Same Day Surgery database	—[c]
Outpatient, primary care, and specialist visits and follow-up long-term care claims	OHIP[d] claims	• MSI physician's billings (MED) • Eligibility group (EGROUP)
Prescription use and follow-up long-term care	Ontario Drug Benefit Claims and Drugs list	• Nova Scotia Drug Information System
Home care services	Home Care Database	• Nova Scotia Health Department of Seniors and Long Term Care
Hospital admissions	DAD[e]	• DAD
Descriptions for international classification of diseases diagnosis codes—version 10 and OHIP diagnosis codes	REF[f]	• REF

[a]PCCF: postal code conversion file.

[b]NACRS: National Ambulatory Care Reporting System.

[c]Not available.

[d]OHIP: Ontario health insurance plan.

[e]DAD: Discharge Abstract Database.

[f]REF: reference files.

Primary Outcome

The primary outcome of interest was to support older adults with complex care needs to safely remain in their home and avoid or delay admission to higher levels of care, such as long-term hospitalization and LTC. The primary outcome was defined as time to a client's inability to return home given the need for LTC or hospital admission under an alternative level of care designation within 1 year of enrollment in the trial. This was ultimately determined by the most responsible provider. The secondary outcome of health service use and mortality was assessed to confirm similar mortality rates within 1 year of enrollment in the trial. Clients were followed up to 1 year until February 28, 2020, to control for the impact of the COVID-19 pandemic. Participants enrolled before February 28, 2019, were included to ensure 1 year of follow-up before the COVID-19 pandemic.

Data Analysis

Baseline characteristics were summarized using descriptive statistics: continuous variables as means and SDs and categorical variables as proportions. Baseline characteristics between the control and intervention participants were examined to verify successful randomization using standardized differences that when greater than 0.1, indicates a potentially meaningful difference [33]. In accordance with ICES and HDNS privacy policies, cell sizes less than or equal to 5 were not reported.

A Cox proportional hazards model was used to evaluate our primary outcome of unable to return home, censoring on mortality and 1 year after enrollment, and our secondary outcome of mortality, only censoring on 1 year after enrollment. Hazard ratios (HRs) with 95% CI were reported. Proportional hazards were assessed using interaction with time terms, and no violations were observed.

Subgroup analyses compared health care use 1 year before and after enrollment for home care clients to determine what effect the intervention had on participants' use of the health care system. For the subgroup analysis, a paired analysis was conducted using the McNemar test for proportions and the Wilcoxon signed rank test for means to compare differences in health care use 1 year before and after enrollment in both patients (Ontario: n=108; Nova Scotia: n=105) and caregivers (Ontario: n=104; Nova Scotia: n=101). All analyses were conducted using SAS (version 9.4; SAS Institute). A statistical significance was defined as a 2-tailed α value less than .05.

Ethical Considerations

This study was approved by the Research Ethics Board at Western University (registration IRB 00000940) and at Nova Scotia Health (registration IRB 102203). All participants

provided written informed consent and could withdraw from the study at any time without penalty. All data were deidentified and assigned a study ID number to support data linkages throughout data collection and analyses. Participants received no compensation to complete the study; intervention group participants received the PRM systems at no personal expense.

Results

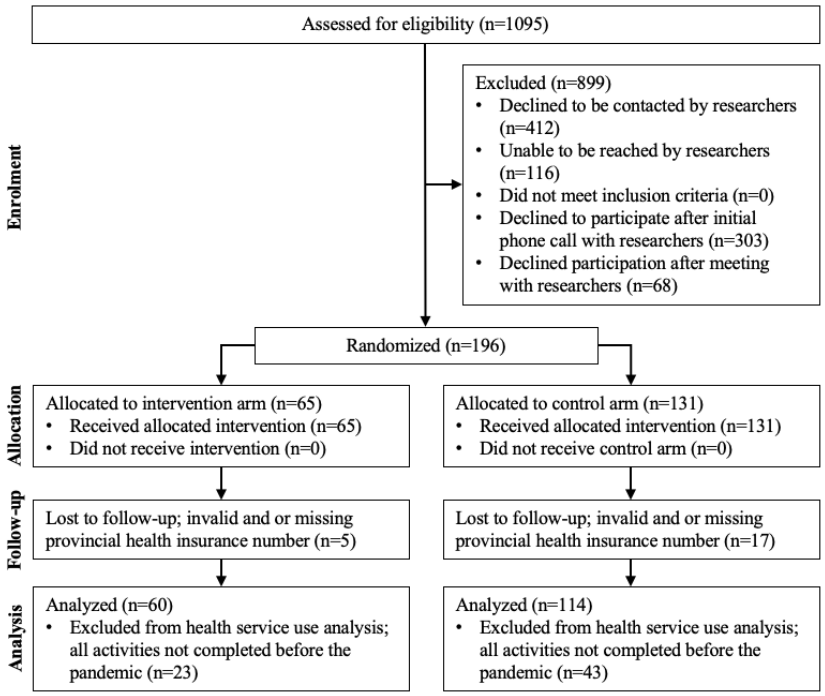
Due to challenges joining administrative patient data—such as data collected as part of this study—across provincial jurisdictions, a single analysis was not possible. The following sections present the results of the separate analyses conducted by ICES in Ontario and HDNS in Nova Scotia to address this study's research question.

Participant Characteristics

Ontario Cohort

The total Ontario cohort comprised 339 participants (home care clients: n=174; caregivers: n=165). In total, 196 home care clients were enrolled in Ontario from April 11, 2017, to January 15, 2020, of which 22 were excluded due to invalid or missing health card numbers, leaving 174 home care clients to be included in analysis. We were able to link 165 caregivers to the home care clients. In Ontario, 116 participants (home care clients: n=60; caregivers: n=56) included in this analysis received the study-provided PRM intervention plus their normal home care and 223 participants (home care clients: n=114; caregivers: n=109) received their normal home care without PRM. Figure 2 provides the randomization and analysis study flow diagram for the Ontario home care client cohort.

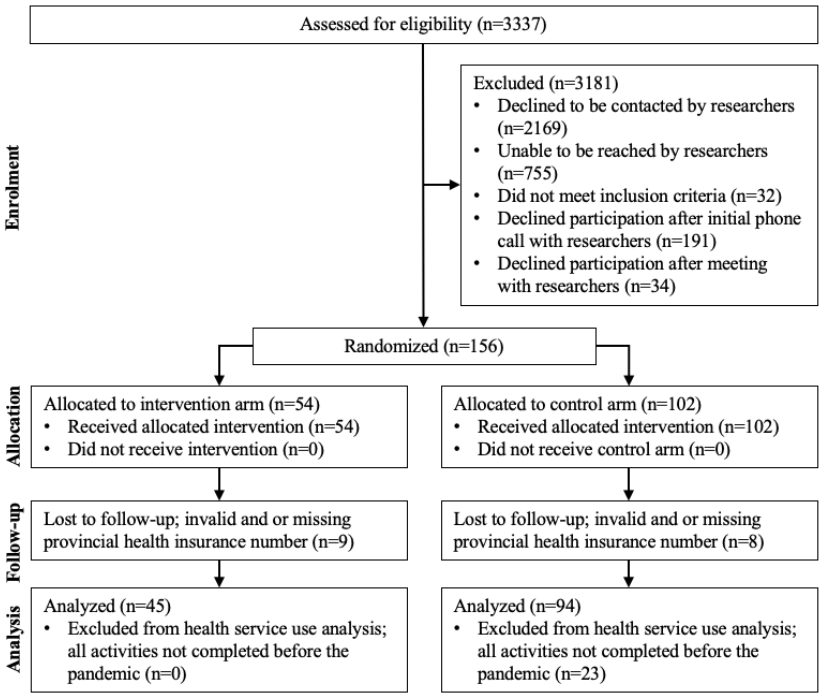
Figure 2. Ontario home care client cohort randomization and analysis flow diagram.



Nova Scotia Cohort

The total Nova Scotia cohort comprised 273 participants (home care clients: n=139; caregivers: n=134). In total, 156 home care clients were enrolled in Nova Scotia from November 29, 2017, to December 16, 2019, of which 17 were excluded due to invalid or missing health card numbers, leaving 139 home care clients to be included in the analysis. We were able to link 134 caregivers to the home care clients. In Nova Scotia, 92 participants (home care clients: n=45; caregivers: n=47) included in this analysis received the study-provided PRM intervention plus their normal home care and 181 participants (home care clients: n=94; caregivers: n=87) received their normal home care without PRM. Figure 3 provides the randomization and analysis study flow diagram for the Nova Scotia home care client cohort.

Figure 3. Nova Scotia home care client cohort randomization and analysis flow diagram.



Home Care Clients

Ontario Home Care Clients

In Ontario, there were no significant differences between home care clients in intervention and control groups at baseline on any of the demographic characteristics evaluated (all $P>.05$). At baseline, Ontario home care clients were on average aged 78 (SD 6.9) years. There was a higher proportion of male participants in the intervention group (40/60, 67%) compared with the control group (57/114, 50%). Most home care clients (125/174, 73%) resided in midsize urban centers. Approximately 20% (35/174) of home care clients resided in the lowest-income neighborhoods (quintile 1), while about 15% (26/174) resided in the highest-income neighborhoods. Most Ontario home care clients (109/174, 63%) lived with their spouse. Nearly all Ontario home care clients (>106/174, >90%) were at medium or high risk of hospitalization. Ontario home care clients had a variety of chronic disease comorbidities, most notably hypertension (144/174, approximately 83%), diabetes (95/174, approximately 55%), and dementia (67/174, approximately 40%). Complete home care client baseline characteristics are presented in Table 2.

Table 2. Ontario and Nova Scotia home care client baseline characteristics.

Characteristics	Ontario		Nova Scotia	
	Control (n=114)	Intervention (n=60)	Control (n=94)	Intervention (n=45)
Age (y), mean (SD)	77.9 (6.9)	78.1 (7.0)	80.1 (7.9)	81.5 (7.3)
Sex, n (%)				
Female	57 (50)	20 (33.3)	62 (66)	33 (73.3)
Male	57 (50)	40 (66.7)	32 (34)	12 (26.7)
Level of rurality, n (%)				
Large CMA[a] (≥500,000)	0 (0)	0 (0)	0 (0)	£5[b]
Midsize CMA (100,000-500,000)	80 (70.2)	45 (75)	66 (70.2)	35 (77.8)
Small CMA (10,000-100,000)	24 (21.1)	9 (15)	13 (13.8)	6 (13.3)
Non-CMA	10 (8.8)	6 (10)	15 (16)	£5
Neighborhood-level income quintile, n (%)				
Quintile 1 or missing	24 (21.1)	14 (23.3)	16 (17)	9 (20)
Quintile 2	20 (17.5)	15 (25)	29 (30.9)	17 (37.8)
Quintile 3	28 (24.6)	9 (15)	19 (20.2)	6 (13.3)
Quintile 4	23 (20.2)	10 (16.7)	13 (13.8)	7 (15.6)
Quintile 5	14 (12.3)	12 (20)	17 (18.1)	6 (13.3)
Living arrangements, n (%)				
Alone	≤5	0 (0)	56 (59.6)	29 (64.4)
With spouse	71 (62.3)	38 (63.3)	12 (12.8)	≤5
With children	0 (0)	0 (0)	0 (0)	0 (0)
Other	≤45	22 (36.7)	0 (0)	≤5
Missing	0 (0)	0 (0)	26 (27.6)	8 (18)
HARP[c], n (%)				
Low risk	≤5	≤5	8 (8.5)	≤5
Medium risk	73 (64)	33 (55)	43 (45.7)	≤20
High risk	≤40	≤30	40 (42.6)	23 (51.1)
Comorbidities (prevalence in sample), n (%)				
Hypertension	95 (83.3)	49 (81.7)	31 (33)	21 (46.7)
Diabetes	61 (53.5)	34 (56.7)	29 (30.9)	12 (26.7)
Dementia	42 (36.8)	25 (41.7)	31 (33)	12 (26.7)
COPD[d]	40 (35.1)	18 (30)	15 (16)	10 (22.2)
CHF[e]	30 (26.3)	16 (26.7)	7 (7.4)	6 (13.3)
Asthma	25 (21.9)	10 (16.7)	≤5	≤5
AMI[f]	6 (5.3)	≤5	≤5	≤5

[a]CMA: census metropolitan areas.

[b]Small cell sizes (ie, n£5) suppressed following the Institute for Clinical Evaluative Sciences and Health Data Nova Scotia reporting guidelines for maintaining participant confidentiality.

[c]HARP: Hospital Admission Risk Profile.

[d]COPD: chronic obstructive pulmonary disease.

[e]CHF: congestive heart failure.

[f]AMI: acute myocardial infarction.

Nova Scotia Home Care Clients

In Nova Scotia, there were no significant differences between home care clients in intervention and control groups at baseline on any of the demographic characteristics evaluated (all *P*>.05). In Nova Scotia, home care clients were on average 81 (SD 7.5) years of age. At baseline, there was a higher proportion of male participants in the control group (32/94, 34%) compared with the intervention group (12/45, 27%). Most Nova Scotia home care clients (101/139, 75%) resided in midsize urban centers. Approximately 20% (25/139) resided in the lowest-income neighborhoods, while approximately 15% (23/139) resided in the highest-income neighborhoods. The majority of Nova Scotia home care clients (85/139, approximately 60%) lived alone. Nearly all home care clients (>114/139, >90%) were at a medium or high risk of hospitalization. The Nova Scotia home care clients also had a variety of chronic disease comorbidities, most notably hypertension (52/139, approximately 40%), dementia (43/139, approximately 30%), and diabetes (41/139, approximately 30%). Complete home care client baseline characteristics are presented in Table 2.

Family and Friend Caregivers

Ontario Caregivers

In Ontario, there were no significant differences between intervention and control caregivers at baseline for any of the characteristics evaluated (all *P*>.05). Ontario caregivers were on average 67 (SD 12.4) years of age. There was a higher proportion of male participants in the control group (39/109, 36%) compared with the intervention group (13/56, 23%). Most caregivers (116/165, 70%) resided in midsize urban centers. Approximately 20% (32/165) resided in the lowest-income neighborhoods, while about 17.6% (29/165) resided in the highest-income neighborhoods. About 40% (<55/165) of Ontario caregivers attained a postsecondary education. Just over half of the caregivers in both groups (99/165, 60%) were retired, with 10% (12/109) of control and 25% (15/56) of intervention caregivers reporting full-time employment. Approximately 66% (114/165) of caregivers were participating in the study with their spouse or partner, and over 80% (144/165) of caregivers lived with the home care client they supported. Approximately 33% (58/165) indicated they would consider moving the home care client to a LTC facility. Ontario caregivers also had a variety of chronic disease comorbidities, most notably hypertension (control: 56/109, 51%; intervention: 37/56, 66%), diabetes (32/165, 20%), and chronic obstructive pulmonary disease (29/165, 18%). Ontario caregivers provided an average of 109 to 119 hours of care or support each week. Scores on the Zarit Burden Interview suggest that caregivers in both groups were similarly burdened by caregiving responsibilities and that such levels of caregiver burden approached high levels of burden. Complete caregiver baseline characteristics are presented in Table 3.

Table 3. Ontario and Nova Scotia family and friend caregiver baseline characteristics.

Characteristics	Ontario		Nova Scotia	
	Control (n=109)	Intervention (n=56)	Control (n=87)	Intervention (n=47)
Age (y), mean (SD)	66.8 (12.2)	67.3 (12.6)	65.4 (10.9)	60.6 (12.1)
Hours spent providing care or support for home care client per week, mean (SD)	109.0 (68.8)	119.0 (67.9)	80.9 (69.3)	80.6 (71.4)
Caregiver burden[a], mean (SD)	17.8 (9.7)	16.8 (8.4)	17.0 (9.3)	18.3 (11.7)
Sex, n (%)[b]				
Female	70 (64.2)	43 (76.8)	55 (63.2)	33 (70.2)
Male	39 (35.8)	13 (23.2)	32 (36.8)	14 (29.8)
Level of rurality, n (%)[b]				
Large CMA[c] (≥500,000)	≤5	0 (0)	0 (0)	0 (0)
Midsize CMA (100,000-500,000)	75 (68.8)	41 (73.2)	61 (70.1)	38 (80.9)
Small CMA (10,000-100,000)	23 (21.1)	8 (14.3)	13 (14.9)	6 (12.8)
Non-CMA	≤17	7 (12.5)	13 (14.9)	≤5
Missing	≤5	0 (0)	0 (0)	≤5
Neighborhood-level income quintile, n (%)[b]				
Quintile 1	22 (20.2)	10 (17.9)	15 (17.2)	6 (12.8)
Quintile 2	18 (16.5)	19 (33.9)	20 (23)	12 (25.5)
Quintile 3	≤40	7 (12.5)	16 (18.4)	13 (27.7)
Quintile 4	24 (22)	7 (12.5)	15 (17.2)	7 (14.9)
Quintile 5	16 (14.7)	13 (23.2)	21 (24.1)	9 (19.1)
Missing	≤5	0 (0)	0 (0)	0 (0)
Self-reported highest level of education, n (%)[b]				
Elementary school	21 (18.4)	9 (15)	10 (11.5)	7 (14.9)
High school	38 (33.3)	17 (28.3)	25 (28.7)	18 (38.3)
College, CEGEP[d], or university	40 (35.1)	28 (46.7)	42 (48.3)	15 (31.9)
Postgraduate degree	12 (10.5)	≤5	10 (11.5)	≤5
Other, declined, or N/A[e]	≤5	≤5	0 (0)	≤5
Missing	≤5	≤5	0 (0)	0 (0)
Self-reported employment status, n (%)[b]				
Full time	12 (10.5)	15 (25)	23 (26.4)	13 (27.7)
Part time	20 (17.5)	≤5	10 (11.5)	9 (19.1)
Leave of absence	≤5	0 (0)	≤5	≤5
Retired	65 (57)	34 (56.7)	48 (55.2)	14 (29.8)
Not employed	14 (12.3)	6 (10)	≤5	10 (21.3)
Declined or N/A	0 (0)	0 (0)	0 (0)	≤5
Missing	≤5	≤5	≤5	≤5
Relationship to home care client, n (%)[b]				
Spouse or partner	73 (64)	41 (68.3)	34 (39.1)	14 (29.8)
Adult child	25 (21.9)	16 (26.7)	35 (40.2)	24 (51.1)
Parent	6 (5.3)	0 (0)	≤5	≤5

Characteristics	Ontario		Nova Scotia	
	Control (n=109)	Intervention (n=56)	Control (n=87)	Intervention (n=47)
Other, declined, or N/A	≤15	≤5	17 (19.5)	8 (17)
Missing	≤5	≤5	≤5	≤5
Living with home care client, n (%[b])				
Yes	95 (83.3)	49 (81.7)	59 (67.8)	30 (63.8)
Would consider moving home care client to long-term care facility, n (%)[b]				
No	≤80	≤50	64 (73.6)	32 (68.1)
Yes	38 (33.3)	20 (33.3)	22 (25.3)	14 (29.8)
Declined or N/A	≤5	0 (0)	≤5	≤5
Missing	≤5	≤5	≤5	≤5
Comorbidities, n (%)[b]				
Hypertension	56 (51.4)	37 (66.1)	34 (39.1)	13 (27.7)
Diabetes	19 (17.4)	13 (23.2)	22 (25.3)	7 (14.9)
COPD[f]	17 (15.6)	12 (21.4)	≤5	≤5
CHF[g]	6 (5.5)	≤5	0 (0)	0 (0)
Asthma	10 (9.2)	7 (12.5)	≤5	≤5
Dementia	≤5	0 (0)	≤5	0 (0)
AMI[h]	≤5	≤5	≤5	0 (0)

[a]Caregiver burden assessed using the 12-item Zarit Burden Interview [29]. Scores range from 0 to 48, with higher scores reflecting higher levels of caregiver burden and scores of 17 and higher indicating a high level of caregiver burden.

[b]Small cell sizes (ie, n£5) suppressed following the Institute for Clinical Evaluative Sciences and Health Data Nova Scotia reporting guidelines for maintaining participant confidentiality.

[c]CMA: census metropolitan area.

[d]CGEP: collège d'enseignement général et professionnel.

[e]N/A: not applicable.

[f]COPD: Chronic Obstructive Pulmonary Disease.

[g]CHF: congestive heart failure.

[h]AMI: acute myocardial infarction.

Nova Scotia Caregivers

In Nova Scotia, there were no significant differences between intervention and control caregivers at baseline for any of the characteristics evaluated (all *P*>.05). On average, caregivers were aged 63 (SD 11.6) years. Approximately 35% (46/134) of Nova Scotia caregivers were male. Most caregivers (99/134, 75%) resided in midsize urban centers. Approximately 15% (21/134) of Nova Scotia caregivers resided in the lowest-income neighborhoods, while about 20% (30/134) resided in the highest-income neighborhoods. Postsecondary school was the highest level of education for 48% (42/87) of control caregivers and 32% (15/47) of intervention caregivers. Less than 30% (36/134) of Nova Scotia caregivers were employed full time; approximately 50% (48/87) of control caregivers and 30% (14/47) of intervention caregivers were retired. Approximately 35% (48/134) were participating with their spouse or partner, while approximately 45% (59/134) were participating with their parent. Approximately 66% (99/134) of Nova Scotia caregivers lived with the home care client they supported. Approximately 25% (36/134) of caregivers indicated they would consider moving the home care client to an LTC facility. Nova Scotia caregivers also had a variety of chronic disease comorbidities, most notably hypertension (control: 34/87, 39%; intervention: 13/47, 28%) and diabetes (control: 22/87, 25%; intervention: 7/47, 15%). Nova Scotia caregivers provided an average of approximately 81 hours of care or support each week. Scores on the Zarit Burden Interview suggests that caregivers in both groups were similarly burdened by caregiving responsibilities and that such levels of caregiver burden constituted high levels of burden. Complete caregiver baseline characteristics are presented in Table 3.

Home Care Clients' Admission to Higher Levels of Care

For the primary outcome in Ontario, there was a median follow-up time of 365 days, and the intervention group had an event rate of 5.5 per 10,000 person-days compared with the event rate of 8.0 per 10,000 person-days for the control group (Table 4). The intervention group was associated with a nonsignificant 30% reduction in risk of being unable to return home (HR 0.7, 95% CI 0.3-1.4). For the secondary outcome in

Ontario, home care client mortality, there was a median follow-up time of 365 days, and the intervention group had an event rate of 2.9 per 10,000 person-days vs the control group that had an event rate of 2.5 per 10,000 person-days (Table 4). Both groups had similar risk of mortality (HR 1.2, 95% CI 0.4-3.2).

Table 4. Unadjusted hazard ratio of primary and secondary outcomes and 95% CIs.

	Ontario			Nova Scotia		
	Control (n=114)	Intervention (n=60)	*P* value	Control (n=94)	Intervention (n=34)	*P* value
Primary outcome: unable to return home						
Events, n (%)	27 (24)	10 (17)	—[a]	8 (9)	≤5	—
Follow-up in days, median (IQR)	365 (257-365)	365 (334-365)	—	365 (204-365)	365 (200-365)	—
Event rate per 10,000 person-days	8.0	5.5	—	2.9	3.2	—
Values, hazard ratio (95% CI)	Reference	0.7 (0.3-1.4)	.30	Reference	1.1 (0.3-3.7)	.90
Secondary outcome: mortality						
Events, n (%)	10 (9)	6 (10)	—	16 (17)	8 (18)	—
Follow-up in days, median (IQR)	365 (365-365)	365 (365-365)	—	365 (211-365)	365 (214-365)	—
Event rate per 10,000 person-days	2.5	2.9	—	5.8	6.2	—
Values, hazard ratio (95% CI)	Reference	1.2 (0.4-3.2)	.80	Reference	1.1 (0.5-2.5)	.90

[a]Not applicable.

For the primary outcome in Nova Scotia, there was a median follow-up of 365 days, and the intervention group had an event rate of 3.2 per 10,000 person-days compared with the event rate of 2.9 per 10,000 person-days for the control group (Table 4). The intervention group was associated with a nonsignificant 10% increase in risk of being unable to return home (HR 1.1, 95% CI 0.3-3.7). For the secondary outcome in Nova Scotia, home care client mortality, there was a median follow-up time of 365 days, and the intervention group had an event rate of 6.2 per 10,000 person-days versus the control group that had an event rate of 5.8 per 10,000 person-days (Table 4). Both groups had similar risk of mortality (HR 1.1, 95% CI 0.5-2.5).

Health Care Use

Ontario Home Care Clients

There was no significant difference in health care use at the 1-year poststudy enrollment for Ontario control home care clients for all services included in the analysis (all $P>.05$). For Ontario intervention home care clients, significantly fewer participants used emergency departments ($P=.01$) at the 1-year poststudy enrollment (24/37, 65%) compared with prestudy enrollment (31/37, 84%). In addition, on average, Ontario intervention home care clients made significantly fewer visits to outpatient services (enrollment: mean 42.2, SD 31.0; after enrollment: mean 32.9, SD 26.0; $P=.04$) and specialist services (enrollment: mean 20.7, SD 25.3; after enrollment: mean 14.6 SD 18.4; $P=.02$). There were no differences in hospital admissions, primary care visits, or prescriptions at the 1-year poststudy enrollment for Ontario intervention home care clients. Complete health care use outcomes for both groups of Ontario home care clients are presented in Table 5.

Table 5. Ontario home care client health care use at the prestudy and 1-year poststudy enrollment.

Health care use	Control			Intervention		
	Before enrollment (n=71)	After enrollment (n=71)	*P* value	Before enrollment (n=37)	After enrollment (n=37)	*P* value
Hospital admissions						
Values, n (%)	26 (36.6)	31 (43.7)	.35	21 (56.8)	14 (37.8)	.07
Values, mean per client (SD)	0.7 (1.2)	0.7 (1.1)	.72	0.9 (1.1)	0.7 (1.0)	.38
Emergency department visits						
Values, n (%)	50 (70.4)	47 (66.2)	.56	31 (83.8)	24 (64.9)	.01
Values, mean per client (SD)	2.1 (2.4)	2.3 (2.8)	.99	2.5 (2.6)	2.0 (2.2)	.39
Outpatient physician visits						
Values, n (%)	71 (100)	71 (100)	—[a]	37 (100)	37 (100)	—
Values, mean per client (SD)	37.5 (35.3)	32.1 (28.2)	.35	42.2 (31.0)	32.9 (26.0)	.04
Primary care visits						
Values, n (%)	71 (100)	66-71 (92.9-100)	NR[b]	37 (100)	37 (100)	NR
Values, mean per client (SD)	14.6 (16.9)	13.0 (11.6)	.86	13.3 (9.5)	12.6 (10.8)	.45
Specialist visits						
Values, n (%)	66-71 (92.9-100)	64 (90.1)	NR	32-37 (86.5-100)	32-37 (86.5-100)	NR
Values, mean per client (SD)	16.0 (23.4)	13.0 (16.5)	.25	20.7 (25.3)	14.6 (18.4)	.02
Prescriptions						
Values, n (%)	66-71 (92.9-100)	66-71 (92.9-100)	NR	32-37 (86.5-100)	32-37 (86.5-100)	NR
Values, mean per client (SD)	6.5 (5.3)	6.7 (5.2)	.27	6.4 (4.5)	5.9 (4.4)	.24

[a]Not applicable.

[b]NR: not reported.

Nova Scotia Home Care Clients

On average, Nova Scotia home care clients in the control group made significantly fewer visits to primary care at the 1-year poststudy enrollment (enrollment: mean 9.2, SD 5.7; after enrollment: mean 7.8, SD 5.7; *P*=.004). There was no significant difference in health care use at the 1-year poststudy enrollment for Nova Scotia control home care clients for all other services included for the analysis (all *P*>.05). At the 1-year poststudy enrollment, Nova Scotia intervention home care clients averaged significantly more hospitalization visits (enrollment: mean 0.3, SD 0.5; after enrollment: mean 0.6, SD 0.9; *P*=.04) and emergency department visits (enrollment: mean 0.9, SD 1.3; after enrollment: mean 1.5, SD 1.8; *P*=.02) and significantly fewer outpatient services (enrollment: mean 16.6, SD 13.2); after enrollment: mean 13.1, SD 11.4; *P*=.005) and primary care visits (enrollment: 8.5, SD 6.8; after enrollment: mean 6.7, SD 5.5; *P*=.02). There were no differences in prescriptions at the 1-year poststudy enrollment for Nova Scotia intervention home care clients (all *P*>.05). Complete health care use outcomes for both groups of Nova Scotia home care clients are presented in Table 6.

Table 6. Nova Scotia home care client health care use at the prestudy and 1-year poststudy enrollment.

Health care use	Control			Intervention		
	Before enrollment (n=71)	After enrollment (=71)	*P* value	Before enrollment (n=34)	After enrollment (n=34)	*P* value
Hospital admissions						
Values, n (%)	19 (26.8)	27 (38)	.10	9 (26.5)	15 (44.1)	.08
Values, mean per client (SD)	0.3 (0.6)	0.5 (0.8)	.07	0.3 (0.5)	0.6 (0.9)	.04
Emergency department visits						
Values, n (%)	34 (47.9)	39 (54.9)	.28	18 (52.9)	22 (64.7)	.21
Values, mean per client (SD)	1.2 (1.7)	1.1 (1.5)	.82	0.9 (1.3)	1.5 (1.8)	.02
Outpatient physician visits						
Values, n (%)	71 (100)	71 (100)	—[a]	32 (94.1)	34 (100)	—
Values, mean per client (SD)	17.4 (9.9)	16.9 (17.2)	.03	16.6 (13.2)	13.1 (11.4)	.005
Primary care visits						
Values, n (%)	71 (100)	69 (97.2)	—	32 (94.1)	33 (97.1)	.56
Values, mean per client (SD)	9.2 (5.7)	7.8 (5.7)	.004	8.5 (6.8)	6.7 (5.5)	.02
Specialist visits						
Values, n (%)	65 (91.5)	58 (81.7)	.07	26 (76.5)	27 (79.4)	.65
Values, mean per client (SD)	5.0 (5.7)	5.6 (14.1)	.54	4.4 (5.3)	3.8 (4.1)	.39
Prescriptions						
Values, n (%)	63 (88.7)	65 (91.5)	.53	29 (85.3)	32 (94.1)	.18
Values, mean per client (SD)	5.3 (3.8)	5.3 (3.8)	.63	5.3 (4.3)	5.3 (3.9)	.62

[a]Not available.

Discussion

Primary Outcome of PRM Within Home Care

While our study did not yield statistically significant results regarding the effectiveness of the PRM model in prolonging home stays, the observed trends suggest that technology-assisted aging in place may be a valuable goal for older adults. Despite the underpowered design of our study, the use of the PRM model of care was associated with meaningful, albeit nonstatistically significant, differences between control and intervention groups on the primary outcome (ie, home care clients' ability to stay at home) in Ontario and Nova Scotia. Specifically, our findings with the Ontario intervention participants trended toward a greater number of home care clients staying home and staying home longer compared with those without PRM in their home. This is similar to the findings of others who have examined the impact of remote monitoring technologies on patient and caregiver well-being [34,35]. The Canadian Expert Panel on Aging highlights the importance of even insignificant delays to admission to LTC [2]. They report that retaining and appropriately supporting older adults at home—even for a month—positively impacts the overall LTC system incapacity or capacity and growing waitlist for care. Furthermore, Canadian older adults' overwhelmingly prefer to age at home and 11% to 30% of Canadians admitted to LTC could have remained at home and in their communities if adequate home care and community supports were available [36,37].

It is important to highlight several participant characteristics that contextualize the nature of this study's patient-caregiver dyads when considering participants' ability to stay at home. The PRM intervention in this study purposefully targeted a select cohort of older adult home care clients (mean age range 77-81 years), many (125/314, 40%) with limited personal resources (ie, education or income) but all who required significant supportive care to remain in their home environment. Home care clients in Ontario were mostly cared for by a spouse living in the same home residence, whereas in Nova Scotia, most older adults lived alone. All older adult home care clients in this study were dealing with significant comorbidities; the most prevalent were hypertension, diabetes, and dementia. In Ontario, about 2 in 5 home care clients were managing dementia, as were approximately 1 in 3 of those in Nova Scotia. This is significant, as previous research demonstrates that older adults (ie, home care clients in this study), especially those with

dementia, express a greater quality of life and social connections, less loneliness, and fewer depressive symptoms when living at home compared with institutionalized LTC [38-41]. The health status of home care clients in this study was characteristic of older adults at risk for higher levels of care, with dementia being the greatest predictor of LTC home admissions [2].

Secondary Outcomes of PRM Within Home Care

What was notable among Ontario intervention home care clients was a significant reduction in their health service use over the course of the study. We found that compared with those receiving usual care, a smaller proportion of Ontario older adult home care clients in the intervention group used emergency departments after 1 year compared with baseline; no such difference was observed in home care clients in the control arm. In addition, Ontario intervention home care clients experienced a significant decrease in the average number of visits to outpatient and specialist services after 1 year, while there were no changes in average visits to these services for home care clients in the control arm. This is an important finding given the *at-risk* health status of older adult home care recipients in the study. Finch et al [35] reported a reduction in health care costs reflective of fewer hospitalizations and emergency visits among "nursing home–eligible" older adults who were monitored by home-based PRM sensors to track their safety (eg, falls) and activities of daily living (eg, time in bed, toilet use, and opening and closing the refrigerator). Consistent with the findings of this study, Finch et al [35] concluded that PRM systems may contribute to a reduction in older adult health service use and associated reductions in health care costs.

However, this was not the case among home care clients in Nova Scotia who received the intervention. In contrast to their Ontario counterparts, Nova Scotia intervention home care clients averaged significantly greater hospital admissions and emergency department visits but significantly fewer outpatient and primary care visits. Discussion with Nova Scotia project partners (health care decision makers) provided some insight into the unanticipated health service use trends of Nova Scotia home care clients in our study. Consistent with our findings, Nova Scotia health system partners reported an observed trend of increased health system use among the Nova Scotia older adult population. The provincial health care decision makers conveyed that emergency services and hospitalization may be a work-around strategy people use to expedite access to LTC services ("skip the line") and circumvent existing policies governing LTC access. Current policies elevate the LTC priority status of hospitalized at-risk older adults who are on the LTC waitlist and given priority consideration for LTC placement (personal communication, June 2024). Home surveillance of older adults' behaviors, activities, and habits through PRM technologies by caregivers may also have contributed to increased health service use. Caregivers' surveillance of older adults' activities in their home may have illuminated older adults' need for greater care. In addition, PRM may generate greater insight into older adults' care needs that may influence (ie, increase) their health service use. This is consistent with pilot study findings of PRM home care services conducted in Nova Scotia (Report of CareLink Evaluation: The Nova Scotia experience, unpublished data, 2015).

Our findings of increased hospital health service use patterns may also be consequential to the 12% to 17% shortfall of primary care service provision within Nova Scotia [42]. With underresourced primary health care services, it may be that Nova Scotians, including complex care older adults needing health care services, would be inclined to access tertiary care services in the absence of primary care services. The Canadian Expert Panel on Aging cautioned that the lack of adequate home care and community-based health care services creates the conditions for premature institutionalization of older adults into LTC. Concernedly, current estimates suggest that up to 1 in 3 Canadians admitted to an LTC home could have remained home with adequate home care services [2]. Avoiding or delaying institutionalized care eases the health system burden particularly in LTC health settings that are perpetually underresourced and ill-equipped to manage the current and pending need for older adult care [43]. An estimated 52,000 Canadian adults are on waiting lists for a placement into an LTC home, while more than 430,000 are estimated to have unmet home care needs, with about 167,000 of them being aged ≥65 years [6,43].

Qualitative research evaluating active remote monitoring and PRM technologies has supported the use of PRM technology as a way to keep at-risk older adults in their home or extend their time at home before requiring institutionalized care [44]. The study, conducted in Nova Scotia, engaged expert key informants including provincial health care policy makers, home care providers, and remote monitoring technology experts to provide their insight into the use of home-based remote monitoring technologies. Key informant participants supported the use of active and PRM technologies as a way to keep complex care older adults at home or extend their time at home. The mitigated risk of adverse events, greater responsiveness to health emergencies (eg, falls), enhanced independence among older adults, and enhanced oversight and greater safety were reported with home monitoring compared with the oversight of care within LTC facilities [44].

Limitations and Strengths

This study was limited by a restricted sample size included in the primary outcomes analysis. Our study was profoundly affected by the COVID-19 pandemic response; public health measures to mitigate the spread of COVID-19 changed the nature of home care and expectations of family and friend caregivers in the middle of this study's sampling frame. This impacted the nature of the family caregiving relationships and the nature of support that caregivers were able to provide. The pandemic response also impacted what was considered normal standard of care for home care services due to resources being redirected to other sectors within the publicly funded health care systems in Ontario and Nova Scotia. Despite this, we did observe promising trends in the data related to PRM as a potentially supportive home care strategy among a cohort of older adults with significant and complex care needs. In addition, we acknowledge the unique opportunity to collect data via PRM systems (ie, information from the sensor notifications) regarding older adults' behaviors within the home setting and the ability to leverage PRM data into large datasets to be analyzed and used to better understand the home care needs of complex care older adults.

Perhaps a strength and limitation of the PRM model of home care for health care systems is the reliance on unpaid caregivers to receive and act on PRM system notifications and not the health care system. Alternatively, the PRM model of home care enhances unpaid caregivers' responsibility and accountability to the care of their family member.

There was also limited representation of marginalized communities within our study, which limits our understanding of the impact of PRM home care models among culturally diverse populations. Finally, the fact that most caregivers were spouses to the patient participants limited this study's understanding of how PRM technologies may impact caregivers in the sandwich generation (ie, those individuals caring for older adult parents and children dependents).

Future Research

There are several opportunities for future research. An economic analysis of the PRM model of care will provide greater insight regarding the implementation of the model of care among health system decision makers. Additional research is needed to fully explore concerns related to home care client, caregiver, and home care provider privacy. Research is needed into the contributions of the PRM model of home care intervention to resolve the challenges of limited health human resources within home care while balancing the health care (physical, emotional, and financial) services given to family caregivers. We also propose an opportunity for data science research in analyzing the aggregated data on older adults' behaviors to leverage historical tracking of data to support trend analysis and the development of best practices within the home care setting and, importantly, to generate a plan of home care services targeted to the unique needs of complex care older adults. A consideration of this PRM model of care against available DIY or off-the-shelf sensor systems requires investigation. The use of off-the-shelf systems has implications for enhanced digital skills with caregivers and older adults that is not required to the same extent with the PRM model of care in this study. Investigation into older adults' and caregiver digital health literacy skills will further inform uptake and use of the PRM model of home care. Research into system data visualization—dashboard components for older adult clients and caregivers—would also support informed decision-making and provide insight into user-friendly dashboards. Longitudinal studies with populations with different chronic diseases (eg, cancer and dementia) is also needed to determine PRM model impact.

Conclusions

The combination of growing waitlists for LTC, limited home care resources, and older adults' preference to remain in their home as long as possible creates the care context for PRM technologies. These technologies hold promise to support complex care older adults to safely remain in their home and avoid or delay admission to higher levels of care. Home is preferred by most older adults even if their health conditions challenge their ability to live independently; almost all Canadians aged ≥65 years reported their intention to take extraordinary measures to avoid institutionalization in an LTC setting. The PRM model of care in this study relied on unpaid caregivers, not the health care system, to receive and act on PRM notifications. While not statistically significant, the findings of this study demonstrate a trend in favor of PRM to support patients' aging in place and positive changes in health service use among older adult home care recipients. To appreciate the provincial differences reported in this study, the PRM model of care, as with any health care system intervention, must be considered within the larger health, social, and political context. Further study is required to understand if longer follow-up time allows more effects of PRM on patients' avoidance of higher levels of care to be detected.

Acknowledgments

This study was supported by the Institute for Clinical Evaluative Sciences, which is funded by an annual grant from the Ontario Ministry of Health and the Ministry of Long-Term Care. Portions of the data used in this report were made available by Health Data Nova Scotia of Dalhousie University. Although this research analysis is based on data obtained from the Nova Scotia Department of Health and Wellness, the observations and opinions expressed are those of the authors and do not represent those of either Health Data Nova Scotia or the Department of Health and Wellness. This study also received funding from Canadian Institutes of Health Research. This document used data adapted from the Statistics Canada Postal Code Conversion File, which is based on data licensed from Canada Post Corporation, and data adapted from the Ontario Ministry of Health Postal Code Conversion File, which contains data copied under license from Canada Post Corporation and Statistics Canada. The authors thank IQVIA Solutions Canada Inc for use of their Drug Information File. Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information, Ontario Health, and the Ontario Ministry of Health. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred.

Authors' Contributions

The work for this manuscript was led by LD and BH. This study was conceptualized by LD (co-lead), SR (co-lead), GW, LW, and SS. Data curation was done by BH and JR. Formal analysis was completed by JR (lead), MR, and SS. Funding acquired by LD (co-lead), SR (co-lead), GW, LW, and SS. Investigation was coordinated by KL and BH. Data analysis methodology was done by JR, SS, LD, BH, MR, and GW. Project administration was done by KL, LD, GW, ER, and BH. Resources were secured by LD, SR, GW, and LW. Overall study supervision was provided by LD, SR, and BH. Results validation was done by LD, BH,

JR, MR, GW, ER, and SR. Visualization of manuscript was done by BH and LD. Writing of the original draft was done by LD, BH, GW, JR, SS, and MR. Reviewing and editing was done by LD (co-lead), BH (co-lead), GW, MR, JR, SS, ER, SR, LW, and KL.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 472 KB - aging_v8i1e69107_app1.pdf]

References

1. Turcotte M. Insights on Canadian society: Canadians with unmet home care needs. Statistics Canada. 2014 Sep 09. URL: https://www150.statcan.gc.ca/n1/en/pub/75-006-x/2014001/article/14042-eng.pdf?st=c76z_elg [accessed 2024-04-29]
2. National Institute on Ageing. Ageing in the right place: supporting older Canadians to live where they want. National Institute on Ageing. 2022. URL: https://static1.squarespace.com/static/5c2fa7b03917eed9b5a436d8/t/638e0857c959d1 [accessed 2024-08-01]
3. Pandemic perspectives on long-term care: insights from Canadians in light of COVID-19. National Institute on Ageing. 2021. URL: https://digitallibrary.cma.ca/link/digitallibrary5 [accessed 2022-08-01]
4. Seniors in need, caregivers in distress : what are the home care priorities for seniors in Canada? Health Council of Canada. 2012. URL: https://publications.gc.ca/site/eng/423913/publication.html [accessed 2024-04-29]
5. Brining care home. Expert Group on Home and Community Care. 2015. URL: https://www.onecaresupport.ca/Files/Bringing%20Care%20Home.pdf [accessed 2024-04-29]
6. Gilmour H. Unmet home care needs in Canada. Health Reports 2018 Nov 21;29(11):3-11 [FREE Full text] [Medline: 30485384]
7. MacDonald BJ, Wolfson M, Hirdes J. The future cost of long-term care in Canada. National Institute of Ageing. 2019 Oct. URL: https://tinyurl.com/yycuz3fd [accessed 2022-08-01]
8. Williams AP, Peckham A, Watkins J, Warrick N, Tam T, Rudoler D, et al. Caring for caregivers: facing up to tough challenges. Healthc Q 2014 Oct 30;17(3):20-23. [doi: 10.12927/hcq.2014.24015] [Medline: 25735060]
9. Blais R, Sears NA, Doran D, Baker GR, Macdonald M, Mitchell L, et al. Assessing adverse events among home care clients in three Canadian provinces using chart review. BMJ Qual Saf 2013 Dec 04;22(12):989-997 [FREE Full text] [doi: 10.1136/bmjqs-2013-002039] [Medline: 23828878]
10. Doran D, Blais R, Baker GR, Harrison M, Lang A, Macdonald M, et al. The safety at home study: an evidence base for policy and practice change. Healthc Q 2014 Oct 30;17(3):42-47. [doi: 10.12927/hcq.2014.24019] [Medline: 25591609]
11. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2023 Mar 01;42(1):20-32. [doi: 10.1017/S0714980822000198] [Medline: 35912590]
12. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152 [FREE Full text] [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
13. Swartz K, Collins L. Caregiver care. Am Fam Physician 2019 Jun 01;99(11):699-706 [FREE Full text] [Medline: 31150177]
14. Barnard D, Yaffe MJ. What is the physician's responsibility to a patient's family caregiver? Virtual Mentor 2014 May 01;16(5):330-338 [FREE Full text] [doi: 10.1001/virtualmentor.2014.16.05.ecas1-1405] [Medline: 24847702]
15. Atlantic Institute on Aging Carelink Advantage Pilot. Fredricton, NB: Orion Marketing Research; 2011.
16. Deis K, Gardner L, Ohinmaa A, Wanke M, Wilson P. Summative Evaluation of the Continuing Care Technology Innovation Pilot Project (CCTI). Edmonton, AB: Charis Management Consulting Inc; 2012.
17. Donelle L, Regan S, Kerr M, Zwarenstein M, Bauer M, Warner G, et al. Caring near and far by connecting community-based clients and family member/friend caregivers using passive remote monitoring: protocol for a pragmatic randomized controlled trial. JMIR Res Protoc 2020 Jan 10;9(1):e15027 [FREE Full text] [doi: 10.2196/15027] [Medline: 31922492]
18. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability. Trials 2009 Jun 03;10(1):37 [FREE Full text] [doi: 10.1186/1745-6215-10-37] [Medline: 19493350]
19. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337(sep29 1):a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
20. Husereau D, Jacobs P, Manns B, Hoomans T, Marshall D, Tamblyn R. Economic evaluation of complex health system interventions: a discussion paper. Institute of Health Economics. 2014. URL: https://www.ihe.ca/publications/economic-evaluation-of-complex-health-system-interventions-a-discussion-paper [accessed 2024-04-29]
21. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015 May 08;350:h2147 [FREE Full text] [doi: 10.1136/bmj.h2147] [Medline: 25956159]

22. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Kim J, Shin W. How to do random allocation (randomization). Clin Orthop Surg 2014 Mar;6(1):103-109 [FREE Full text] [doi: 10.4055/cios.2014.6.1.103] [Medline: 24605197]
24. Bowling A, Hankins M, Windle G, Bilotta C, Grant R. A short measure of quality of life in older age: the performance of the brief Older People's Quality of Life questionnaire (OPQOL-brief). Arch Gerontol Geriatr 2013 Jan;56(1):181-187 [FREE Full text] [doi: 10.1016/j.archger.2012.08.012] [Medline: 22999305]
25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
26. Sager MA, Rudberg MA, Muhammad Jalaluddin MS, Franke T, Inouye SK, Landefeld CS, et al. Hospital admission risk profile (HARP): identifying older patients at risk for functional decline following acute medical illness and hospitalization. J Am Geriatr Soc 1996 Mar 27;44(3):251-257. [doi: 10.1111/j.1532-5415.1996.tb00910.x] [Medline: 8600192]
27. Selected caregiver assessment measures: a resource inventory for practitioners. Family Caregiver Alliance & Benjamin Rose Institute on Aging. 2012. URL: https://www.carealliance.ie/userfiles/file/SelCGAssmtMeas_ResInv_FINAL_12_10_12.pdf [accessed 2024-04-29]
28. California caregiver resource centers uniform assessment tool. California Caregiver Resource Centers. 2003. URL: https://www.caregiver.org/caregivers-count-too-s4-assessment-tool-examples [accessed 2024-04-29]
29. Bédard M, Molloy D, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
30. Tarlow BJ, Wisniewski SR, Belle SH, Rubert M, Ory MG, Gallagher-Thompson D. Positive aspects of caregiving: contributions of the REACH project to the development of new measures for Alzheimer's caregiving. Res Aging 2004 Jul 01;26(4):429-453. [doi: 10.1177/0164027504264493]
31. Koopman C, Pelletier KR, Murray JF, Sharda CE, Berger ML, Turpin RS, et al. Stanford presenteeism scale: health status and employee productivity. J Occup Environ Med 2002 Jan;44(1):14-20. [doi: 10.1097/00043764-200201000-00004] [Medline: 11802460]
32. Todtman K, Gustafson AW. The financial impact scale: an instrument for assessing long-term caregivers. J Gerontol Soc Work 2008 Oct 25;18(1-2):135-150. [doi: 10.1300/j083v18n03_10]
33. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. P T 2008 Dec;33(12):700-711 [FREE Full text] [Medline: 19750051]
34. Gaugler JE, Zmora R, Mitchell LL, Finlay J, Rosebush CE, Nkimbeng M, et al. Remote activity monitoring for family caregivers of persons living with dementia: a mixed methods, randomized controlled evaluation. BMC Geriatr 2021 Dec 18;21(1):715 [FREE Full text] [doi: 10.1186/s12877-021-02634-8] [Medline: 34922475]
35. Finch M, Griffin K, Pacala JT. Reduced healthcare use and apparent savings with passive home monitoring technology: a pilot study. J Am Geriatr Soc 2017 Jun;65(6):1301-1305. [doi: 10.1111/jgs.14892] [Medline: 28407212]
36. 1 in 9 new long-term care residents potentially could have been cared for at home. Canadian Institute for Health Information. 2020. URL: https://www.cihi.ca/en/1-in-9-new-long-term-care-residents-potentially-could-have-been-cared-for-at-home [accessed 2022-08-01]
37. Seniors in transition: exploring pathways across the care continuum. Canadian Institute for Health Information. 2017. URL: https://www.cihi.ca/en/seniors-in-transition-exploring-pathways-across-the-care-continuum [accessed 2022-08-01]
38. Nikmat AW, Hawthorne G, Al-Mashoor SH. The comparison of quality of life among people with mild dementia in nursing home and home care--a preliminary report. Dementia (London) 2015 Jan 08;14(1):114-125. [doi: 10.1177/1471301213494509] [Medline: 24339093]
39. Olsen C, Pedersen I, Bergland A, Enders-Slegers MJ, Jøranson N, Calogiuri G, et al. Differences in quality of life in home-dwelling persons and nursing home residents with dementia - a cross-sectional study. BMC Geriatr 2016 Jul 11;16(1):137 [FREE Full text] [doi: 10.1186/s12877-016-0312-4] [Medline: 27400744]
40. Ramocha LM, Louw QA, Tshabalala MD. Quality of life and physical activity among older adults living in institutions compared to the community. S Afr J Physiother 2017;73(1):342 [FREE Full text] [doi: 10.4102/sajp.v73i1.342] [Medline: 30135899]
41. Mahmood A, Yamamoto T, Lee M, Steggell C. Perceptions and use of gerotechnology: implications for aging in place. J Hous Elderly 2008 Jun 16;22(1-2):104-126. [doi: 10.1080/02763890802097144]
42. Li K, Frumkin A, Bi WG, Magrill J, Newton C. Biopsy of Canada's family physician shortage. Fam Med Community Health 2023 May 12;11(2):e002236 [FREE Full text] [doi: 10.1136/fmch-2023-002236] [Medline: 37173094]
43. Cost estimate for Motion 77: improvements to longterm care. Office of the Parliamentary Budget Officer. URL: https://publications.gc.ca/site/eng/9.903146/publication.html [accessed 2022-08-01]
44. Warner G, Weeks LE, Chen YT, Hiebert B, Ledoux K, Donelle L. Key informant perceptions of challenges and facilitators to implementing passive remote monitoring technology for home care clients. Gerontechnology 2023;22(1):1-13 [FREE Full text] [doi: 10.4017/gt.2023.22.1.834.09]

Abbreviations

HDNS: Health Data Nova Scotia
HR: hazard ratio
ICES: Institute for Clinical Evaluative Sciences
LTC: long-term care
PRCT: pragmatic randomized controlled trial
PRM: passive remote monitoring

Edited by H Köttl; submitted 23.11.24; peer-reviewed by S Ofori, S Mukherjee; comments to author 22.12.24; revised version received 28.01.25; accepted 14.02.25; published 19.03.25.

Please cite as:
Donelle L, Hiebert B, Warner G, Reid M, Reid J, Shariff S, Richard E, Regan S, Weeks L, Ledoux K
Passive Remote Monitoring Technologies' Influence on Home Care Clients' Ability to Stay Home: Multiprovincial Randomized Controlled Trial
JMIR Aging 2025;8:e69107
URL: https://aging.jmir.org/2025/1/e69107
doi:10.2196/69107
PMID:

Exploring Smart Health Wearable Adoption Among Singaporean Older Adults Based on Self-Determination Theory: Web-Based Survey Study

Hyunjin Kang[1], PhD; Tingting Yang[1], MA; Nazira Banu[1], MA; Sheryl Wei Ting Ng[1], MA; Jeong Kyu Lee[2], PhD

[1]Wee Kim Wee School of Communication and Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

[2]Department of Health and Exercise Science, University of Oklahoma, Norman, OK, United States

Corresponding Author:
Hyunjin Kang, PhD
Wee Kim Wee School of Communication and Information, Nanyang Technological University, 31 Nanyang Link, Singapore, Singapore

Abstract

Background: Smart health wearables offer significant benefits for older adults, enabling seamless health monitoring and personalized suggestions based on real-time data. Promoting adoption and sustained use among older adults is essential to empower autonomous health management, leading to better health outcomes, improved quality of life, and reduced strain on health care systems.

Objective: This study investigates how autonomy-related contextual factors, including artificial intelligence (AI) anxiety, perceived privacy risks, and health consciousness, are related to older adults' psychological needs of competence, autonomy, and relatedness (RQ1). We then examined whether the fulfillment of these needs positively predicts older adults' intentions to adopt these devices (H1), and how they mediate the relationship between these factors and older adults' intentions to use smart health wearables (RQ2). Additionally, it compares experienced and nonexperienced older adult users regarding the influence of these psychological needs on use intentions (RQ3).

Methods: A web-based survey was conducted with individuals aged 60 years and above in Singapore, using a Qualtrics survey panel. A total of 306 participants (177 male; mean age of 65.47 years, age range 60 - 85 years) completed the survey. A structural equation model was used to analyze associations among AI anxiety, perceived privacy risks, and health consciousness, and the mediating factors of competence, autonomy, and relatedness, as well as their relationship to smart health wearable use intention.

Results: Health consciousness positively influenced all intrinsic motivation factors—competence, autonomy, and relatedness—while perceived privacy risks negatively affected all three. AI anxiety was negatively associated with competence only. Both privacy risk perceptions and health consciousness were indirectly linked to older adults' intentions to use smart health wearables through competence and relatedness. No significant differences were found in motivational structures between older adults with prior experience and those without.

Conclusions: This study contributes to the application of self-determination theory in promoting the use of smart technology for health management among older adults. The results highlight the critical role of intrinsic motivation—particularly competence—in older adults' adoption of smart health wearables. While privacy concerns diminish motivation, health consciousness fosters it. The study results offer valuable implications for designing technologies that align with older adults' motivations, potentially benefiting aging populations in other technologically advanced societies. Developers should focus on intuitive design, transparent privacy practices, and social features to encourage adoption, empowering older adults to use smart wearables for proactive health management.

(JMIR Aging 2025;8:e69008) doi:10.2196/69008

KEYWORDS

smart health wearables; self-determination theory; AI anxiety; perceived privacy risk; health consciousness

Introduction

Background

Smart health wearables are devices powered by artificial intelligence (AI) designed to be worn on the body, enhancing health management. This technology holds significant promise for the older population, offering a nonintrusive way to continuously monitor health information and provide valuable assistance [1,2]. The use of smart health wearables has been verified in clinical intervention research and proven to assist in self-health management and prevention of chronic diseases,

such as diabetes, hypertension, and dementia [3]. With technological advancements today, some of these technologies have already been integrated into personal devices, such as smartphones and other mobile devices, making it easier for older adults to track their health conveniently [4]. To fully leverage the potential of health wearables, it is essential to promote the adoption and sustained use among older adults, empowering them in autonomous health management. This can lead to better health outcomes, an improved quality of life, and reduced strain on the health care system [5].

The promotion of adoption and sustained use of smart health wearables can be understood through the lens of self-determination theory (SDT) [6,7], which posits that fostering intrinsic motivation, rather than extrinsic, is essential for encouraging prolonged active engagement in a behavior. SDT suggests that fostering an environment which supports an individual's self-determination in engaging in a specific behavior is crucial for satisfying fundamental human needs that drive intrinsic motivation [6]. We examine AI anxiety, privacy risk perceptions, and health consciousness as the critical contextual factors shaping self-determination of using smart health wearables. These 3 contextual factors are closely tied to key smart health wearable aspects. (1) AI anxiety [8] stems from the automated, AI-driven recommendations in smart health wearables, which can make users feel they are losing control over health decisions. (2) Privacy risk perceptions [9] relate to the continuous data collection by smart health wearables; if inadequately protected, users' sense of control over their personal information can be reduced. (3) Health consciousness aligns with smart health wearables' purpose of empowering users to take charge of their health, reinforcing autonomy for proactive self-management [10].

Another key proposition of SDT is that fulfilling the 3 psychological needs of competence, relatedness, and autonomy is essential for intrinsic motivation [6]. When applied to smart health wearables for older adults, the extent to which these devices help older adults feel capable, enable independent choices, and foster a sense of connection with others can shape their intrinsic motivation to engage with the devices and promote sustained usage.

Taken together, this study explores how 3 contextual factors related to self-determination in using smart health wearables—AI anxiety, privacy risk perceptions, and health consciousness—are associated with fulfilling the needs for competence, relatedness, and autonomy, thereby promoting use intentions. Additionally, the study examines whether differences exist between experienced and nonexperienced wearable users in terms of the psychological needs influencing their intentions. Understanding these distinctions provides insight into the motivators for each group at different adoption stages, offering evidence for targeted strategies to sustain user engagement among older adults and thereby promote continuous, effective self-health management [11].

Through this study, we aim to contribute to the literature on self-determination in health promotion through smart technology by identifying key factors that intrinsically motivate older adults to use smart health wearables, which ultimately enhances their long-term well-being. A web-based survey has been conducted in Singapore, one of the world's most connected nations with high information and communication technology penetration and a robust AI ecosystem [12]. Furthermore, Singapore has recently become a "super-aged society," with nearly one-fifth of its population over 65 years old [13]. As such, the insights gained from this research can offer valuable, generalizable implications for designing technologies that align with older adults' motivations, potentially benefiting aging populations in other technologically advanced societies.

Literature Review: Smart Health Wearables for Older Adults' Self-Health Management

Wearables refer to "intelligent computing devices integrated into various accessories, including clothing, fashion accessories, and other everyday items worn by consumers" [14]. Smart health wearables are wearable devices designed to enhance users' health management, offering users a seamless and integrated user experience. Smart health wearables have many benefits, with their main benefit being empowering older adults to take charge of their health [15]. By integrating sensors, (eg, bio, motion, and environmental sensors), internet connection, as well as AI, smart health wearables enable individual users to track and exchange data, and even make smart and personalized decisions for their health management [1]. Some examples of smart health wearables include biosensors, GPS, and radio frequency identification (RFID) technologies that can monitor health conditions by tracking physical information such as blood pressure, oxygen level, and sleep patterns to provide personalized health feedback and medical suggestions. This feedback is valuable for those managing chronic conditions, such as diabetes or cardiovascular diseases [16]. New and improved wearable designs (eg, smartwatches) incorporate advanced algorithms and machine learning to analyze users' data and provide personalized workout programs and nutrition recommendations [17].

The potential of health wearables to improve self-health management is increasingly recognized. However, adoption rates among older adults remain low, largely due to skepticism about the tangible benefits of using wearables [18,19]. Despite a growing body of research exploring enablers and barriers to adopting smart health wearables, the literature remains fragmented [20]. Further studies are needed to understand how these devices are perceived and what motivates older adults to adopt them long-term. For example, recent studies have reported that older adults are more likely to adopt technology if they perceive a clear and immediate benefit from its usage. If the advantages of wearables are not clearly communicated or experienced, older adults may not see them as worth the investment [21]. However, such insights only scratch the surface. More in-depth exploration is necessary to uncover intrinsically motivated factors that can help older adults perceive a clear benefit from adopting wearables. Understanding these factors is crucial to fostering long-term engagement with these wearable devices. This study draws upon SDT and the motivational technology model to identify intrinsic motivators that can encourage older adults to adopt smart health wearable devices for their self-health management.

Three Contextual Factors Related to Autonomous Motivation

Overview

One of the key propositions of SDT is that creating an environment that supports one's self-determination in engaging in a given behavior is essential for fulfilling the key human needs required for fostering intrinsic motivation. In environments that respect and encourage individuals' autonomy, they are more likely to develop a sense of their true selves and achieve self-determination for their actions [6]. In social conditions that pressure people's autonomous behavior, autonomy is deprived. In such situations, the fulfillment of the 3 needs for intrinsic motivation will be hampered. However, the key human needs for intrinsic motivation will be nurtured in autonomy-supportive contexts, where the external environment supports one's self-initiation and choice.

The role of autonomous motivation in encouraging behavior is particularly significant in health-specific contexts. Being in an autonomy-supportive health care climate, for instance, and having autonomous motivation are linked to successful physical activity and dietary management, which, in turn, leads to better health outcomes for patients with type 2 diabetes [22]. Likewise, feelings of autonomous motivation positively impacted the behavioral intentions and usage behaviors of sports apps among students, contrary to feelings of controlled motivation [23]. However, solely using smart health wearables is insufficient to encourage self-health management. A study that sought to promote physical activity among adolescents using health wearables found that, contrary to expectations, feelings of autonomous motivation decreased over an 8-week period because of feelings of peer comparison that created a social environment where participants' autonomy was undermined [24]. This highlights the importance of contextual factors in shaping autonomous motivation to use smart health wearables for health management. Moreover, the technology acceptance model (TAM) [25] and the unified theory of acceptance and use of technology (UTAUT) [26] explain how users accept a given technology. These theories suggest that external factors, such as social influence, should be considered to understand people's attitudes and intentions to use technology.

This study examines the roles of 3 social-contextual factors—AI anxiety, perceived privacy risks, and health consciousness—that are crucial for creating a context that fosters the autonomy of older adults in engaging in self-health management using smart health wearables.

AI Anxiety

There is growing tension between human agency and machine agency in the context of the rise of AI, especially when user experience relies on the algorithms of smart technology [27]. Deci and Ryan [28] defined human agency as motivated behaviors that emanate from one's integrated self. In the context of communication with smart technology, user agency is defined as "the degree to which the self feels like that he or she is a relevant actor" and a feeling that the technology provides "manipulability" for the user to exercise their influence throughout the interaction [29, p. 61]. The feeling of being in control is essential as it bolsters people's intrinsic motivation.

However, as AI techniques advance at an unprecedented speed, people may begin to experience negative psychological tension driven by their natural desire to limit AI agency [27]. AI anxiety describes the psychological status of fear and trepidation when people are concerned that they will lose autonomy and be controlled by AI [8]. It is caused by the constantly evolving technological advancement of AI, along with the challenges and uncertainty that AI brings [30]. For example, Airbnb hosts experienced this algorithm-related AI anxiety as they navigated a complex ecosystem filled with uncertainty and frustration, since the algorithms comprise both "known and unknown factors" [31, p. 9].

While existing research indicates the prevalence of AI anxiety across various populations [32], AI anxiety may be particularly prominent among older adult users. Research shows that age is positively associated with greater anxiety over computers [33]. Although there are ontological differences between traditional computing technologies and AI, AI technologies may generate a wider range of anxieties because of their complexity and perceived autonomy [34]. Furthermore, low levels of technological literacy among older adults may worsen AI anxiety since a key dimension of AI literacy involves the ability to learn AI, which can be particularly challenging for older adults [34,35].

AI anxiety becomes even more pronounced when AI is used to deliver health care services, like in the case of smart health wearables. The decision-making processes of AI models that directly affect the user's well-being may be beyond the understanding of users [36]. Studies have highlighted the fear and frustration expressed by older adults when using smart health wearables to monitor their blood oxygen saturation levels [37], which illustrates how smart health wearables can deepen pre-existing anxieties. Since studies indicate that technological anxiety can negatively affect attitudes and intentions to use technology [3,38,39], we investigate AI anxiety as a significant factor shaping intrinsic motivation for using smart health wearables.

Perceived Privacy Risks

Privacy risks, throughout the various stages of designing, distributing, and using smart health wearables, can cause privacy violations. Technically, (1) most wearable devices are equipped with various sensors (eg, motion, location, and physiological sensors) that continuously collect data, (2) these data are highly personal and thus sensitive, and (3) opportunistic data usage can lead to privacy breaches [40]. Hence, using smart health wearables for health management inevitably triggers people's concerns about smart health wearables' ability to protect privacy. Complicated terms regarding data collection and sharing further exacerbate this concern, with many users unaware of how their data are managed [18,41]. Additionally, incidents involving unauthorized access and cyberattacks compound these concerns [40]. Particularly for older adult users, concerns about privacy and data security are heightened [9].

Moreover, privacy is essential for developing an autonomous self. It is commonly accepted that the human body and health data are "the central mediator of autonomy and individual privacy" [42, p. xvi]. Therefore, individuals' controllability over their information [43] or "selective control of access to the self or to one's group" [44, p. 18] is the core element of privacy. Through such processes of boundary regulation, individuals decide how open they want to be to others. However, smart health wearables need health-related personal information to function (eg, customize exercise plans and make smart recommendations). Hence, there exists this tension between (1) the need to grant controlling power of private information to smart health wearables and (2) people's need for autonomous decision-making about information disclosure. For example, when using smart home applications integrating sensors and other network technologies, older adult users expressed their concern about their controllability over their privacy; however, this sense of tension can be ameliorated if they have "full control of the sensor," meaning they would feel more in control if they have a say about when to activate or deactivate the device [45, p. 4750]. Thus, perceived privacy risks are closely tied to feelings of autonomy, and the control of personal information is key to intrinsic motivation for using health care wearables [46].

Health Consciousness

Health consciousness refers to how individuals care about their health and the level at which they are motivated to engage in preventive health practices such as health information seeking, home-based exercise, and healthy food choices [47]. Specifically, health consciousness reflects people's orientation toward three health-related dimensions: (1) personal health awareness, (2) self-responsibility, and (3) health motivation [48]. Health consciousness has repeatedly been shown to have a positive effect on health-related behaviors like using the internet to search for health information [10], choosing foods based on their health benefits [47], or using dietary and fitness apps [49].

There are several reasons to explain why health consciousness motivates older adults to engage in self-health management. Health-conscious people tend to have a stronger belief in their ability to control their health, which influences their decisions to engage in health-promoting activities [50]. Furthermore, health consciousness can motivate people to be proactive in looking for ways to evaluate their conditions, like taking up information and communication technology to mitigate loneliness [51]. Proactiveness is a distinct characteristic of health consciousness and differs from simply reacting to negative health outcomes. This aligns with research using the health belief model, which suggests that health-related internet use was more strongly proactively motivated by health consciousness, rather than being influenced by perceptions of one's health risk in a reactive manner [10]. Smart health wearables that are designed to enable self-directed health management activities [52] precisely meet the needs of health-conscious individuals. Thus, it would be fruitful to assess health consciousness as a potential factor encouraging the adoption of smart health wearables among older adults.

Three Basic Needs for Self-Determination

We used the SDT [6,7] to examine the factors related to the motivation for adopting smart health wearables among older adults. The core idea of SDT is that intrinsic motivation is necessary for sustaining engagement in a particular behavior. The 3 basic needs—competence, relatedness, and autonomy—are fundamental sources of intrinsic motivation and the natural tendency for self-development.

Competence refers to the ability to achieve personal goals. Autonomy is the freedom to make self-initiated choices, and relatedness reflects feeling connected and cared for by social groups or communities. When individuals perceive that a particular behavior fulfills these 3 basic needs, they tend to find it enjoyable (ie, intrinsically motivating) and will engage in it without relying on external rewards. Hence, from the lens of SDT, this study examines how the 3 psychological needs shape intrinsic motivation and ensures the long-term use of smart health wearables for older adults' self-health management.

As with other interactive technologies, design elements in smart health wearables can influence the sense of competence, autonomy, and relatedness through their intelligent, user-focused features [53]. For example, features such as goal setting and reward badges can enhance users' competence by providing feedback on their achievements and progress. Yet, older adults often struggle with setting up or understanding the data generated by these devices, which can undermine their sense of competence [54]. Additionally, while personalized nudges and recommendations from these devices may sometimes compromise users' sense of autonomy by making decisions based on their data [27], customizable interfaces can empower users to feel more in control. Furthermore, social features—such as communities, friend invitations, and participation in challenges—can enhance a sense of relatedness among users [24,53]. Wu and Lim [21] found that older adults are more likely to use wearables when family members or friends in their close social circle also adopt the devices.

Overall, this study explores 3 social-contextual factors influencing the autonomous use of smart health wearables for health management—AI anxiety, perceived privacy risks, and health consciousness—and their potential associations with the psychological needs related to intrinsic motivation among older adults. Accordingly, we pose the following research question:

RQ1: How are autonomy-related contextual factors regarding smart health wearable use (ie, AI anxiety, perceived privacy risks, and health consciousness) associated with the sense of competence, autonomy, and relatedness among older adults?

In addition, we aim to unpack the relationship between the psychological needs of older adults and their intention to use smart health wearables. Self-determined motivation shaped by the fulfillment of these 3 needs plays a crucial role in people's intention to use technology [55]. For instance, a study found that intentions to use health technology, such as wellness clouds for health tracking, were positively associated with intrinsic motivation, driven by expectations of performance, playfulness, and ease of use [56]. Thus, we hypothesize:

H1: Perceived competence, autonomy, and relatedness in using smart health wearables will be positively associated with older adults' intentions to use these devices.

We expect that the needs of competence, autonomy, and relatedness fulfilled by health wearables use would mediate the relationships between the autonomy-related contextual factors and use intention. This study will explore these indirect associations by proposing the following research question:

RQ2: How do perceived competence, autonomy, and relatedness in using smart health wearables mediate the relationships between the autonomous-related contextual factors (ie, AI anxiety, perceived privacy risks, and health consciousness) and older adults' intentions to use these devices?

Experienced Versus Nonexperienced Users

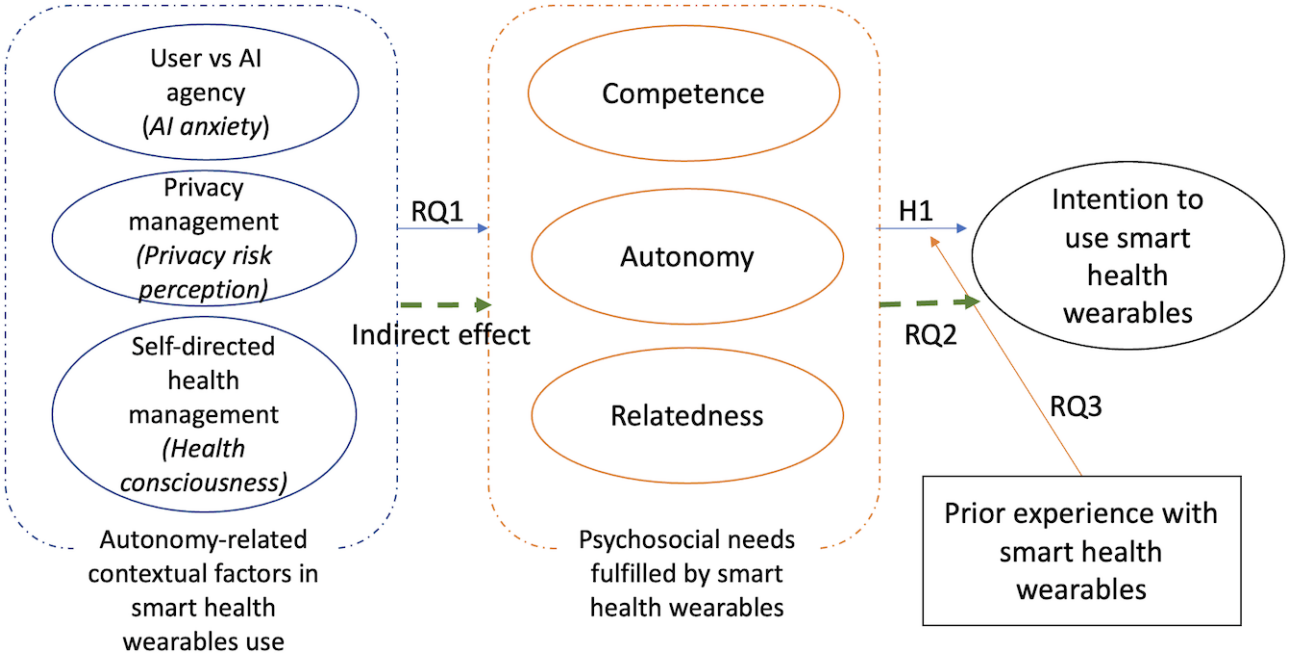
We also examine potential differences in motivation between older adults with and without prior experience using smart health wearables. According to the UTAUT and TAM models, as discussed earlier, users' prior experience with using the technology is also one of the key variables, as prior experience can work as a conditional factor in changing the degree of effect of other antecedent factors on technology acceptance (eg, intention to use and attitudes toward a technology) [26]. Specifically, whether a user has experience using a technology may affect their evaluation of motivations, effort expectancy, social influence, and perceived ease of use, among other factors [26,57]. For example, one study about smart health wearables found that the relation between behavior intention and actual use is stronger for more experienced users, compared with the new users; in addition, the experienced users rated the facilitating conditions (eg, resource and knowledge they received) as less important [58]. Interestingly, among older adults, whether users have prior experience also showed a difference in what types of smart health wearables they choose: compared with experienced users whose choice focuses mostly on smartwatches and wristbands, nonexperienced users' choices were more dispersed, showing a wide variety of preferences such as smart rings and clips among others. This indicates that prior experience might affect users' overall expectations among many technology acceptance indicators [59], supporting the proposition that prior experience may be a conditional factor that differentiates users' motivational structure predicting their technology adoption intentions [60], as experienced users tend to exhibit greater confidence and skill in navigating wearable features [61,62]. Experienced older adults may more readily recognize benefits such as improvements in health metrics or physical activity [16]. Conversely, older adults with limited or no prior experience may have different expectations regarding wearables, often accompanied by uncertainty. While such uncertainty can foster positive expectations about the benefits of devices, it may also provoke feelings of intimidation. Older adults without prior experience may struggle to fully grasp the benefits of smart health wearables early on. Moreover, without clear evidence of the device's impact on their well-being, these older adults might question the use of the technology, resulting in lower intrinsic motivation to adopt such devices [63].

However, the moderating effect of prior experience from existing literature shows inconsistent effects. For example, one study examined whether users' prior experience of using robots predicts their approval of autonomous delivery robots, and the correlation was not significant [60]. In light of these discrepancies, we will examine whether and how these 2 groups—those with and without prior experience—differ in their motivational structures for using smart health wearables in health management. Understanding the motivational differences between those who are experienced versus not experienced users of smart health wearables offers valuable insights into the distinct needs of users in different adoption stages. These insights can guide the development of more user-friendly designs and tailored educational strategies, ultimately encouraging wider adoption among hesitant or inexperienced individuals. Accordingly, we propose the following research question:

RQ3: How are experienced older adult users different from the nonexperienced older adult group regarding the relationship between the 3 psychological needs and use intention?

Figure 1 summarizes the model tested in this study.

Figure 1. Hypothesized model summary. AI: artificial intelligence; H: hypothesis; RQ: research question.



Methods

Data Collection

A web-based survey was conducted with individuals over the age of 60 years in Singapore from March to April 2022. For data collection, we used a survey panel from Qualtrics, one of the major global survey companies, which provides panels across various age groups in different countries. In total, 306 participants (177 male; mean age of 65.47 years, age range 60 - 85 years; ethnicity: 84.6% [259/306] Chinese, 3.3% [10/306] Malay, 6.5% [20/306] Indian, and 5.6% [17/306] others) completed the survey. At the beginning of the study, they were given descriptions and examples of smart health wearables. Then, they were asked if they had used such devices before. Of the 306 participants, 163 (53.3%) responded "yes" indicating their prior experiences with smart health wearables, and fitness trackers were the most common type of wearable used (135/306, 44.1%). Meanwhile, 143 (46.7%) responded "no" to the question asking about their prior experience.

Ethical Considerations

The survey procedure and materials were reviewed and approved by the institutional review board at Nanyang Technological University (IRB-2022-163). Informed consent was obtained before participants began the survey, and they had the option to opt out. All collected data were deidentified. Compensation was provided to participants by the survey company in the form of credit points in accordance with its policy.

Measurement

AI Anxiety

To measure the anxiety caused by uncertainty about how AI works, 3 items were adapted from Meuter et al [64], originally developed for measuring user anxiety about service technology. A 7-point Likert scale was used (1=strongly disagree to 7=strongly agree; eg, "I feel apprehensive about using AI technology").

Perceived Privacy Risk

Perceived privacy risk regarding wearable health technology was measured using 3 items [65], each rated on a 7-point Likert scale (1=strongly disagree to 7=strongly agree; eg, "It would be risky to disclose my personal health information to wearable device vendors").

Health Consciousness

Health consciousness was measured using a 3-item scale, with participants indicating their agreement with each statement on a 7-point Likert scale (1=strongly disagree to 7=strongly agree; eg, "I take responsibility for the state of my health") [66].

Self-Determination Factors

Adapted from Ryan et al [67], 3 factors of self-determination—competence, autonomy, and relatedness—were measured using a 7-point Likert scale (1=strongly disagree to 7=strongly agree). Competence in using wearable health devices was measured using a 3-item scale (eg, "I would feel competent in using a smart wearable health care device"). Autonomy was assessed with a 4-item scale (eg, "I would feel free to decide for myself how to do things on a smart wearable health care device"). Relatedness to other users of wearable health care devices was measured using 3 items (eg, "I would find the relationships with other users of smart wearable health care devices fulfilling").

Future Use Intention

Intention to adopt smart wearable health care technology was measured using a 3-item semantic differential scale. Participants rated their likelihood, probability, and willingness to use smart wearable health care technology in the next 3 months, with responses ranging from 1 to 7 (eg, likelihood: 1=unlikely to 7=likely).

Refer to Table 1 for the measurement items, internal reliability (Cronbach α), and basic statistics for each factor's composite score.

Table . Measurement items, internal reliability, and descriptive statistics.

Measures and items			Cronbach α	Mean (SD)
AI[a] anxiety			0.78	3.95 (1.28)
	I feel apprehensive about using AI technology.			
	I hesitate to use AI technology for fear of making mistakes I cannot correct.			
	I have difficulty understanding AI-related technological matters.			
Perceived privacy risk			0.92	4.40 (1.44)
	It would be risky to disclose my personal health information to wearable device vendors.			
	There would be a high potential for loss associated with disclosing my personal health information to vendors providing wearable devices.			
	There would be too much uncertainty associated with giving my personal health information to vendors providing wearable devices.			
Health consciousness			0.63	5.37 (1.06)
	I am concerned about my health all the time.			
	I notice how I feel physically as I go through the day.			
	I take responsibility for the state of my health.			
Competence			0.95	4.83 (1.31)
	I would feel competent in using a smart wearable health care device.			
	I would feel capable when using a smart wearable health care device.			
	I would feel like I am effective when using a smart wearable health care device.			
Autonomy			0.91	4.70 (1.17)
	I would feel free to decide for myself how to do things on a smart wearable health care device.			
	I would generally feel free to express my ideas and opinions on a smart wearable health care device.			
	I would feel like I can pretty much be myself when I use a smart wearable health care device.			
	I would experience a lot of freedom when I use a smart wearable health care device.			
Relatedness			0.70	4.03 (.97)
	I would find the relationships with other users of smart wearable health care devices fulfilling.			
	I would find the relationships with other users of smart wearable health care devices important.			
	I would not feel close to other smart wearable health care device users (reverse-coded).			
Intention			0.94	4.78 (1.93)
	How likely are you to use smart wearable health care devices in the next 3 months?			
		Unlikely or likely (1 to 7 scale)		
		Not probable or probable (1 to 7 scale)		

Measures and items	Cronbach α	Mean (SD)
Unwilling or willing (1 to 7 scale)		

[a]AI: artificial intelligence.

Measurement Model

The measurement model was tested using comparative fit index (CFI) analysis. Covariances for the first 3 autonomy items and the first 2 competence items were allowed to improve model fit. After this modification, the model showed good fit: χ^2_{185}=365.9, *P*<.001, root mean square error of approximation (RMSEA)=0.057 (90% CI 0.048-0.065), CFI=0.966, Tucker-Lewis index (TLI)=0.957, and standardized root mean square residual (SRMR)=0.048, all within acceptable ranges.

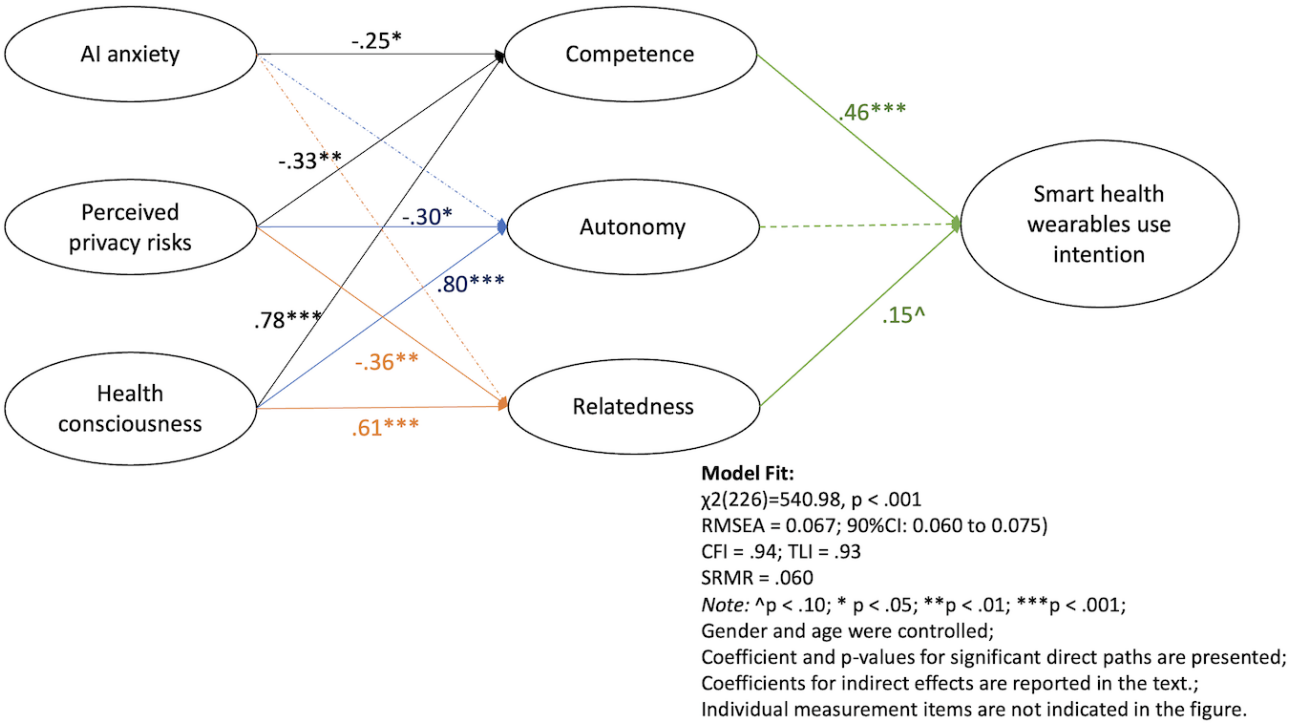
Results

The structural equation model (SEM) results indicate significant associations among AI anxiety, perceived privacy risks, and health consciousness, and the mediating factors of competence, autonomy, and relatedness, which are related to smart health wearable use intention. Model fit indices suggested a good fit: χ^2_{226}=541, *P*<.001; RMSEA=0.067 (90% CI 0.06-0.075); CFI=0.94; TLI=0.93; SRMR=0.06.

For RQ1, AI anxiety was negatively associated only with competence (β=−.25, *P*=.039); perceived privacy risks showed a significant negative association with competence (β=−.33, *P*=.002), autonomy (β=−.3, *P*=.012), and relatedness (β=−.36, *P*=.001). Health consciousness was positively associated with competence (β=.78, *P*<.001), autonomy (β=.8, *P*<.001), and relatedness (β=.61, *P*<.001).

H1 examined the associations between the 3 self-determination model dimensions and the intention to use smart health wearables. Competence was significantly associated with use intention (β=.46, *P*<.001), and relatedness also showed a marginally significant association (β=.15, *P*=.05). Figure 2 summarizes the SEM results.

Figure 2. Structural equation model results for RQ1 and H1. AI: artificial intelligence; CFI: comparative fit index; H: hypothesis; RMSEA: root mean square error of approximation; RQ: research question; SRMR: standardized root mean square residual.



To answer RQ2, we examined the indirect effects of AI anxiety, perceived privacy risks, and health consciousness on use intention through competence, autonomy, and relatedness as mediators, using 5000 bootstrapped samples. The total indirect effect of AI anxiety on use intention was not statistically significant (*B*=−0.31, 95% CI −0.78 to 0.05) although the indirect pathway through competence was significant in a negative direction (*B*=−0.27, 95% CI −0.71 to −0.05). For perceived privacy risks, the total indirect effect was significant in a negative direction (*B*=−0.36, 95% CI −0.59 to −0.14), via competence (*B*=−0.24, 95% CI −0.52 to −0.08) and relatedness (*B*=−0.09, 95% CI −0.23 to −0.004). Lastly, health consciousness showed a significant total indirect effect on use intention in a positive direction (*B*=1.37, 95% CI 1.01-2.08), with a strong positive mediation through competence (*B*=0.97, 95% CI 0.44-1.78) and relatedness (*B*=0.25, 95% CI 0.02-0.58). We found no significant indirect pathways through autonomy for any predictors.

Finally, RQ3 asked if experienced versus not experienced older adult users would have different motivation structures shaping the intention to use smart health wearables. First, measurement

invariance between experienced versus nonexperienced users for the 3 self-determination factors and use intention were tested. The models allowing for free estimation of all parameters across groups showed a good fit to the data: χ^2_{110}=183.8, *P*<.001; RMSEA=0.066, 90% CI 0.049-0.083; CFI=0.979; TLI=0.971; and SRMR=0.045. Additionally, the chi-square difference test comparing the model constraining the factor loadings to be equal across groups, against the unconstrained model, was not significant ($\Delta\chi^2_9$=6.8, *P*=.66), indicating that the factor loadings are invariant across groups. To assess the relationship between the 3 motivation factors and behavioral intention across experienced and nonexperienced user groups, 2 SEMs were estimated: an unconstrained path coefficient model (χ^2_{18}=38.5, *P*=.003, RMSEA=0.086, 90% CI 0.048-0.124, CFI=0.968) and a constrained model (χ^2_{30}=50.5, *P*=.011, RMSEA=0.067, 90% CI 0.032-0.098, CFI=0.968), both showing acceptable model fits. However, the chi-square difference test indicated that the relationships between competence, autonomy, and relatedness, and behavioral intention between the 2 groups do not statistically differ ($\Delta\chi^2_{12}$=12, *P*>.05).

Discussion

Theoretical Implications

Drawing on SDT, this study sheds light on the motivational factors influencing older adults' intention to use smart wearables for health care. By examining external contextual factors that may shape older adults' autonomous use of smart wearables for health management, we provide more nuanced insights into the self-determination model, enhancing the understanding of how to promote self-determined, proactive health management through technology. Moreover, this age-specific insight contributes to a deeper understanding of how SDT can be adapted to different user populations in the smart health wearables context.

Our study examined how external contextual factors influence intrinsic motivation, particularly in the adoption of smart health wearables among older adults (RQ1). The results reveal that perceived privacy risks erode all motivation factors: competence, autonomy, and relatedness. The fear of data breaches or the misuse of personal health information can make older adults feel that they lack control over their personal data, directly undermining their sense of autonomy. Perceived privacy risks also affect competence, as older adults may feel uncertain about their ability to navigate complex privacy settings or how their data are collected and used. This uncertainty creates a barrier to engaging with smart health wearables, as users who feel less competent in managing their personal information may be less likely to adopt such technologies. The negative association between privacy risk perception and relatedness may imply that fears of data misuse or unauthorized access can undermine users' willingness to engage in social features of smart health wearables, such as sharing health information with peers or participating in health communities. We also found a significant association between AI anxiety and competence, among 3 intrinsic motivation factors. AI anxiety is characterized by fear and apprehension about AI's decision-making processes [8]. Older adults, who may already feel less confident in their technological abilities, may find AI's opaque algorithms intimidating, leading them to doubt their capacity to effectively use smart health wearables, consequently reducing their intrinsic motivation to use them.

However, health consciousness strongly supports the fulfillment of all 3 psychological needs among older adult users; health-conscious older adults are more likely to feel competent, autonomous, and socially connected through the use of smart health wearables. This result indicates that the tendency of being proactive in self-health management among health-conscious individuals [47,50] can be largely translated to self-determined smart health wearables adoption. Especially, the significant association between health consciousness and relatedness suggests that older adults who care about their health are more likely to engage with smart health wearables and connect with others through health communities and shared goals, not just for self-health management. These findings suggest that while SDT offers a strong framework for understanding motivation across various domains, it may require adjustments for technology adoption among older adults.

We then examined the associations between the 3 psychological needs that foster intrinsic motivation and the intention to use smart health wearables (H1). The results indicate that, for older adults, feeling capable of using smart health wearables (competence) is a significant driver of wearable use, and experiencing social connection (relatedness) showed a marginally significant association. However, independent decision-making (autonomy) was not a significant factor. This result might be because older adults recognize the benefits of delegating monitoring and decision-making for health management to technology, which provides personalized recommendations for users. This result suggests that older adults may prioritize supported autonomy in using smart health wearables. Studies [29,68] have suggested that users seek to benefit from the autonomous capacity of AI-based technology while maintaining control over it.

It is also worth noting that our result contradicts Jung and Kang [52], who found that only autonomy, among the 3 intrinsic motivation factors, significantly predicted enjoyment in using smart fitness wearables, based on a survey conducted with a general population in the United States. This also indicates a possibility that autonomy might be experienced or expected differently among the older population in Singapore. Future research should explore how older adults balance autonomy and supported autonomy through health technologies and how smart health wearables can foster this balance without diminishing their sense of agency, especially through comparison with young users.

The mediating roles of the 3 psychological needs between contextual factors and the intention to use smart health wearables (RQ2) offer a deeper understanding of the factors influencing self-determination in wearable health technology use. The findings underscore the importance of fostering feelings of competence to encourage smart wearable adoption among older adults. Competence, as a key driver of intrinsic motivation, is particularly crucial for this demographic, who may feel

intimidated or uncertain about using such technology [54]. The strong indirect associations between privacy risk perceptions, health consciousness, and usage intentions through competence suggest that merely addressing privacy concerns and promoting health awareness may be insufficient. Instead, efforts should also focus on clearly communicating privacy safeguards and emphasizing the value of self-directed health management in a way that reinforces older adults' confidence and competence, ultimately helping them integrate wearables into their health routines successfully.

While relatedness played a secondary role, it nonetheless contributes to older adults' intrinsic motivation, aligning with previous research identifying relatedness as an important intrinsic motivator in social media use [69]. Our study indicates that the feelings of being connected to peer users, family, and caregivers remain a significant need for older adult users even when engaging with technology primarily intended for purposes other than social interaction or relationship building.

Finally, the lack of significant differences between experienced and nonexperienced users regarding the associations between motivational factors and use intention (RQ3) suggests that the need for competence and relatedness is important in shaping the intention to adopt smart health wearables, regardless of prior experience with such devices. Previous research found that individuals with prior experience using smart devices (eg, smartphones and smart televisions) perceived them as less complex and were more willing to try new technologies, as compared to those without experience [19]. However, the absence of differences between the 2 groups in our study could be interpreted in 2 ways. First, advancements in user-centered design and AI may have made smart health wearables more intuitive and less cognitively demanding [18], reducing the potential barriers or concerns around competence in using these devices. Alternatively, it is possible that smart health wearables remain complex and intimidating, even for older adults who have had prior experience with similar technology. Future research could use a qualitative approach, such as in-depth interviews with older adult users, to explore the underlying reasons behind this nonsignificant result. In addition, for older adult users, connecting with family and friends was identified as a major gratification in using new media technology [69-71]. Our study demonstrates that, for both experienced and nonexperienced groups, fostering relatedness with others remains an important reason for adopting health wearables among older adults, which often did not hold truth in other populations [24,52].

Practical Implications

Our findings indicate that perceived privacy risks can negatively impact older adults' sense of competence and autonomy when adopting smart health wearables. To address perceived privacy risks, smart health wearables should offer easy-to-use privacy settings that give older adults control over how their data are shared. Transparent communication about data security can enhance users' sense of control and trust, encouraging wider adoption and sustained use.

Competence emerged as the strongest predictor of smart health wearables adoption. Therefore, developers should focus on creating intuitive, easy-to-use devices tailored to older adults, with guides, tutorials, and support resources that enhance users' confidence and ability. Features such as personalized feedback, goal-setting, and progress tracking can further reinforce older adults' sense of accomplishment in managing their health, fostering continued engagement. By helping older adults feel capable and effective, developers can promote long-term use of wearables.

The feeling of social connectedness (relatedness) also plays a significant role in driving smart health wearables adoption. Developers should incorporate social features that allow older adults to connect with family, friends, or health communities—such as shared health data, group challenges, and support groups. These features foster a sense of social support, enhancing motivation. Health care providers and caregivers can encourage older adults to use smart health wearables in social settings, which can strengthen their motivation for ongoing use.

The nonsignificant difference between experienced and nonexperienced users in terms of the motivational structure for smart health wearables use suggests that both groups rely on feeling competent and socially connected when deciding to use smart health wearables. From a practical standpoint, this indicates that interventions aimed at enhancing these motivational factors—such as simplifying the technology to improve competence and fostering social connections—are equally important for both groups. For example, social workers and government organizations can provide training courses for older adults focusing on these aspects. A recent study showed that even short training courses (eg, teaching older adults how to use smartphones for social interaction and medical use) increased the positive effect of older adult users' effort expectancy (how understandable and easy to learn to use the technology) on behavior intention [72]. Therefore, strategies to increase smart health wearables adoption should focus on addressing these core needs universally, rather than differentiating between experienced users and nonexperienced users.

Limitations and Future Research

We acknowledge several limitations in this study. First, our cross-sectional design limits the ability to establish causal relationships between autonomy-related factors and wearables adoption. As a result, the associations observed in this study are correlational. Longitudinal research would be valuable in examining how variables such as AI anxiety, privacy concerns, and health consciousness impact both short- and long-term use of wearables over time, providing a clearer understanding of causality.

Second, this study was conducted via a web-based survey, which inherently requires participants to have a certain level of digital literacy and internet access. Consequently, our sample may not fully represent older adults with limited internet access or low digital literacy. To capture a more inclusive profile of older adult participants, future studies should consider using offline methods, such as paper-based surveys, to ensure a broader representation of older adults with varying digital skills.

Third, while our research primarily focuses on intrinsic motivation among older adult users for adopting smart health wearables, actual behavioral intention (ie, intention to use) is also influenced by external factors, such as the usability designs of the devices. Studies indicate that older adults have specific performance expectations regarding wearable functionality and ease of use [73]. Therefore, future studies should consider both intrinsic and external factors when exploring predictors of older adults' adoption and sustained use of smart health wearables.

Fourth, the high ownership rate of fitness trackers among our participants (135/306, 44.1%) can be attributed to Singapore's National Steps Challenge, initiated in 2015, which provided free fitness trackers to encourage physical activity. This unique context may have influenced our results, highlighting the need for further cross-cultural studies in other countries to gain complementary insights and verify the generalizability of these findings across different cultural and health care contexts.

Finally, according to the National Population and Talent Division, 81.8% of those over 60 years old are ethnic Chinese, 11% are Malay, 8% are Indian, and 1% belong to other ethnicities as of 2024 [74]. While our sample largely reflects the overall ethnic group composition, Chinese (259/306, 84.6%) and other ethnicities (17/306, 5.6%) are slightly overrepresented, and Malay (10/306, 3.3%) and Indian participants (20/306, 6.5%) are slightly underrepresented. Given that cultural and ethnic differences could significantly influence health-related [75] as well as technology and privacy-related factors [76,77], future research should aim for a more representative sample to capture the diverse experiences and preferences of all ethnic groups, ensuring broader applicability of the findings.

Conclusions

Through the lens of SDT, this study advances our understanding of older adults' adoption of smart health wearables. The results highlighted the critical roles of competence, relatedness, and autonomy-supporting contexts in shaping their intrinsic motivation. While perceived privacy risks and AI anxiety negatively impact motivation, health consciousness emerges as a strong enabler of self-determined adoption of health wearables. Our findings emphasize the need for user-friendly designs, robust privacy safeguards, and social features that foster connection and confidence among older adults. By addressing these factors, developers and stakeholders can better support older adults in integrating smart wearables into their proactive health management routines.

Acknowledgments

This study was supported by the Ministry of Education in Singapore Tier 1 Grant (to HK) and the University of Oklahoma Libraries' Open Access Fund (to JKL).

Conflicts of Interest

None declared.

References

1. Hiremath S, Yang G, Mankodiya K. Wearable Internet of Things: concept, architectural components and promises for person-centered healthcare. Presented at: 2014 4th International Conference on Wireless Mobile Communication and Healthcare – Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH); Nov 3-5, 2014; Athens, Greece p. 304-307. [doi: 10.4108/Mobihealth33544.2014.7015971]
2. Peng C, Zhao H, Zhang S. Determinants and cross-national moderators of wearable health tracker adoption: a meta-analysis. Sustainability 2021;13(23):13328. [doi: 10.3390/su132313328]
3. Jeng MY, Pai FY, Yeh TM. Antecedents for older adults' intention to use smart health wearable devices—technology anxiety as a moderator. Behav Sci (Basel) 2022 Apr 18;12(4):114. [doi: 10.3390/bs12040114] [Medline: 35447686]
4. Liu F, Ngai E, Ju X. Understanding mobile health service use: an investigation of routine and emergency use intentions. Int J Inf Manage 2019 Apr;45:107-117. [doi: 10.1016/j.ijinfomgt.2018.09.004]
5. Metcalf D, Milliard STJ, Gomez M, Schwartz M. Wearables and the Internet of Things for health: wearable, interconnected devices promise more efficient and comprehensive health care. IEEE Pulse 2016;7(5):35-39. [doi: 10.1109/MPUL.2016.2592260] [Medline: 28113167]
6. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior: Springer US; 1985. [doi: 10.1007/978-1-4899-2271-7]
7. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
8. Johnson DG, Verdicchio M. AI anxiety. J Assoc Inf Sci Technol 2017 Sep;68(9):2267-2270. [doi: 10.1002/asi.23867]
9. Xing F, Peng G, Zhang B, Li S, Liang X. Socio-technical barriers affecting large-scale deployment of AI-enabled wearable medical devices among the ageing population in China. Technol Forecast Soc Change 2021 May;166:120609. [doi: 10.1016/j.techfore.2021.120609]
10. Ahadzadeh AS, Pahlevan Sharif S, Ong FS, Khong KW. Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J Med Internet Res 2015 Feb 19;17(2):e45. [doi: 10.2196/jmir.3564] [Medline: 25700481]

11. Windasari NA, Lin FR, Kato-Lin YC. Continued use of wearable fitness technology: a value co-creation perspective. Int J Inf Manage 2021 Apr;57:102292. [doi: 10.1016/j.ijinfomgt.2020.102292]
12. AI Singapore. AI Singapore. 2022. URL: https://aisingapore.org [accessed 2024-11-16]
13. Lau D. How Singapore is preparing for a super aged society come 2026. CNA. 2024. URL: https://www.channelnewsasia.com/today/big-read/super-aged-2026-singapore-ready-4656756 [accessed 2024-11-16]
14. Wright R, Keith L. Wearable technology: if the tech fits, wear it. J Electron Resour Med Libr 2014 Oct 2;11(4):204-216. [doi: 10.1080/15424065.2014.969051]
15. Li J, Ma Q, Chan AH, Man SS. Health monitoring through wearable technologies for older adults: smart wearables acceptance model. Appl Ergon 2019 Feb;75:162-169. [doi: 10.1016/j.apergo.2018.10.006] [Medline: 30509522]
16. Cheung ML, Chau KY, Lam MHS, et al. Examining consumers' adoption of wearable healthcare technology: the role of health attributes. Int J Environ Res Public Health 2019 Jun 26;16(13):2257. [doi: 10.3390/ijerph16132257] [Medline: 31247962]
17. Yadav SS, Agarwal R, Bharath K, Rao S, Thakur CS. tinyRadar for fitness: a contactless framework for edge computing. IEEE Trans Biomed Circuits Syst 2023 Apr;17(2):192-201. [doi: 10.1109/TBCAS.2023.3244240] [Medline: 37022890]
18. Garcia Reyes EP, Kelly R, Buchanan G, Waycott J. Understanding older adults' experiences with technologies for health self-management: interview study. JMIR Aging 2023 Mar 21;6:e43197. [doi: 10.2196/43197] [Medline: 36943333]
19. Farivar S, Abouzahra M, Ghasemaghaei M. Wearable device adoption among older adults: a mixed-methods study. Int J Inf Manage 2020 Dec;55:102209. [doi: 10.1016/j.ijinfomgt.2020.102209]
20. Ferreira JJ, Fernandes CI, Rammal HG, Veiga PM. Wearable technology and consumer interaction: a systematic review and research agenda. Comput Human Behav 2021 May;118:106710. [doi: 10.1016/j.chb.2021.106710]
21. Wu C, Lim GG. Investigating older adults users' willingness to adopt wearable devices by integrating the technology acceptance model (UTAUT2) and the Technology Readiness Index theory. Front Public Health 2024;12:1449594. [doi: 10.3389/fpubh.2024.1449594] [Medline: 39421816]
22. Koponen AM, Simonsen N, Suominen S. Determinants of physical activity among patients with type 2 diabetes: the role of perceived autonomy support, autonomous motivation and self-care competence. Psychol Health Med 2017 Mar;22(3):332-344. [doi: 10.1080/13548506.2016.1154179] [Medline: 26952696]
23. Guo J. Influencing factors of college students' use of sports apps in mandatory situations: based on UTAUT and SDT. Biomed Res Int 2022;2022(1):9378860. [doi: 10.1155/2022/9378860] [Medline: 36193303]
24. Kerner C, Goodyear VA. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on Fitbits with adolescents. Am J Health Educ 2017 Sep 3;48(5):287-297. [doi: 10.1080/19325037.2017.1343161]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
27. Sundar SS. Rise of machine agency: a framework for studying the psychology of human–AI interaction (HAII). J Comput Mediat Commun 2020 Mar 23;25(1):74-88. [doi: 10.1093/jcmc/zmz026]
28. Deci EL, Ryan RM. Human autonomy. In: Kernis MH, editor. Efficacy, Agency, and Self-Esteem: Springer US; 1995:31-49. [doi: 10.1007/978-1-4899-1280-0_3]
29. Sundar SS. Self as source: agency and customization in interactive media. In: Mediated Interpersonal Communication: Routledge; 2008:72-88. [doi: 10.4324/9780203926864-12]
30. Lemay DJ, Basnet RB, Doleck T. Fearing the robot apocalypse: correlates of AI anxiety. Int J Learn Anal Artif Intell Educ IJAI 2020;2(2):24. [doi: 10.3991/ijai.v2i2.16759]
31. Jhaver S, Karpfen Y, Antin J. Algorithmic anxiety and coping strategies of Airbnb hosts. Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18); Apr 21-26, 2018; Montreal QC, Canada p. 1-12. [doi: 10.1145/3173574.3173995]
32. Neyazi TA, Ng SWT, Hobbs M, Yue A. Understanding user interactions and perceptions of AI risk in Singapore. Big Data Soc 2023 Jul;10(2). [doi: 10.1177/20539517231213823]
33. Fernández-Ardèvol M, Ivan L. Why age is not that important? An ageing perspective on computer anxiety. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Design for Aging: Springer International Publishing; 2015:189-200. [doi: 10.1007/978-3-319-20892-3_19]
34. Li J, Huang JS. Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. Technol Soc 2020 Nov;63:101410. [doi: 10.1016/j.techsoc.2020.101410]
35. Finkelstein R, Wu Y, Brennan-Ing M. Older adults' experiences with using information and communication technology and tech support services in New York City: findings and recommendations for post-pandemic digital pedagogy for older adults. Front Psychol 2023;14:1129512. [doi: 10.3389/fpsyg.2023.1129512] [Medline: 37138998]
36. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]

37. Arnaert A, Sumbly P, da Costa D, Liu Y, Debe Z, Charbonneau S. Acceptance of the Apple Watch Series 6 for telemonitoring of older adults with chronic obstructive pulmonary disease: qualitative descriptive study part 1. JMIR Aging 2023 Dec 26;6:e41549. [doi: 10.2196/41549] [Medline: 38147371]
38. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
39. Tsai TH, Lin WY, Chang YS, Chang PC, Lee MY. Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. PLoS One 2020;15(1):e0227270. [doi: 10.1371/journal.pone.0227270] [Medline: 31929560]
40. Marakhimov A, Joo J. Consumer adaptation and infusion of wearable devices for healthcare. Comput Human Behav 2017 Nov;76:135-148. [doi: 10.1016/j.chb.2017.07.016]
41. Cilliers L. Wearable devices in healthcare: privacy and information security issues. Health Inf Manag J 2020 May;49(2-3):150-156. [doi: 10.1177/1833358319851684]
42. Akrivopoulou CM. Protecting the Genetic Self From Biometric Threats: Autonomy, Identity, and Genetic Privacy: IGI Global; 2015. [doi: 10.4018/978-1-4666-8153-8]
43. Kupfer J. Privacy, autonomy, and self-concept. Am Philos Q 1987;24:81-89 [FREE Full text]
44. Altman I. The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding: Brooks/Cole Publishing Company; 1975.
45. Townsend D, Knoefel F, Goubran R. Privacy versus autonomy: a tradeoff model for smart home monitoring technologies. Presented at: 2011 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 30 to Sep 3, 2011; Boston, MA. [doi: 10.1109/IEMBS.2011.6091176]
46. Schönau A, Dasgupta I, Brown T, Versalovic E, Klein E, Goering S. Mapping the dimensions of agency. AJOB Neurosci 2021;12(2-3):172-186. [doi: 10.1080/21507740.2021.1896599] [Medline: 33764258]
47. Mai R, Hoffmann S. How to combat the unhealthy = tasty intuition: the influencing role of health consciousness. J Public Policy Mark 2015 Apr;34(1):63-83. [doi: 10.1509/jppm.14.006]
48. Gould SJ. Health consciousness and health behavior: the application of a new health consciousness scale. Am J Prev Med 1990;6(4):228-237. [doi: 10.1016/S0749-3797(18)31009-2] [Medline: 2223170]
49. Chen MF, Lin NP. Incorporation of health consciousness into the technology readiness and acceptance model to predict app download and usage intentions. Internet Res 2018 Apr 4;28(2):351-373. [doi: 10.1108/IntR-03-2017-0099]
50. Pu B, Zhang L, Tang Z, Qiu Y. The relationship between health consciousness and home-based exercise in China during the COVID-19 pandemic. Int J Environ Res Public Health 2020 Aug 6;17(16):5693. [doi: 10.3390/ijerph17165693] [Medline: 32781751]
51. Zhang Q, Guo X, Vogel D. Addressing elderly loneliness with ICT use: the role of ICT self-efficacy and health consciousness. Psychol Health Med 2022 May 28;27(5):1063-1071. [doi: 10.1080/13548506.2020.1847302]
52. Jung EH, Kang H. Self-determination in wearable fitness technology: the moderating effect of age. Int J Human Comput Interact 2022 Sep 14;38(15):1399-1409. [doi: 10.1080/10447318.2021.2002048]
53. Sundar SS, Bellur S, Jia H. Motivational technologies: a theoretical framework for designing preventive health applications. In: Bang M, Ragnemalm EL, editors. Persuasive Technology Design for Health and Safety: Springer; 2012:112-122. [doi: 10.1007/978-3-642-31037-9_10]
54. Lee C, Coughlin JF. PERSPECTIVE: Older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J Prod Innov Manag 2015 Sep;32(5):747-759. [doi: 10.1111/jpim.12176]
55. Rupp MA, Michaelis JR, McConnell DS, Smither JA. The impact of technological trust and self-determined motivation on intentions to use wearable fitness technology. Proc Hum Factors Ergon Soc Annu Meet 2016 Sep;60(1):1434-1438. [doi: 10.1177/1541931213601329]
56. Hsieh PJ. The impact of motivations, health beliefs, and basic human needs on mobile self-management: an extension of the self-determination theory perspective. Behav Inf Technol 2023 Jun 11;42(8):1045-1063. [doi: 10.1080/0144929X.2022.2059007]
57. Lazaro MJS, Lim J, Kim SH, Yun MH. Wearable Technologies: Acceptance Model for Smartwatch Adoption Among Older Adults: Springer International Publishing; 2020:303-315. [doi: 10.1007/978-3-030-50252-2_23]
58. Ozkan-Yildirim S, Pancar T. Smart wearable technology for health tracking: what are the factors that affect their use? In: Marques G, Bhoi AK, Albuquerque V, KS H, editors. IoT in Healthcare and Ambient Assisted Living: Springer; 2021:165-199.
59. Rani N, Chu SL. Wearables can help me learn: a survey of user perception of wearable technologies for learning in everyday life. Educ Inf Technol 2022 Apr;27(3):3381-3401. [doi: 10.1007/s10639-021-10726-6]
60. Xu Z, Li Y, Hao L. An empirical examination of UTAUT model and social network analysis. Library Hi Tech 2022 Feb 14;40(1):18-32. [doi: 10.1108/LHT-11-2018-0175]
61. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7. [doi: 10.2196/mhealth.4225] [Medline: 26818775]

62. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022 Jul 13;10(7):e35684. [doi: 10.2196/35684] [Medline: 35830222]
63. Sun S, Jiang L, Zhou Y. Associations between perceived usefulness and willingness to use smart healthcare devices among Chinese older adults: the multiple mediating effect of technology interactivity and technology anxiety. Digit Health 2024;10. [doi: 10.1177/20552076241254194] [Medline: 38812850]
64. Meuter ML, Ostrom AL, Bitner MJ, Roundtree R. The influence of technology anxiety on consumer use and experiences with self-service technologies. J Bus Res 2003 Nov;56(11):899-906. [doi: 10.1016/S0148-2963(01)00276-4]
65. Li H, Wu J, Gao Y, Shi Y. Examining individuals' adoption of healthcare wearable devices: an empirical study from privacy calculus perspective. Int J Med Inform 2016 Apr;88:8-17. [doi: 10.1016/j.ijmedinf.2015.12.010]
66. Shin DH, Biocca F. Health experience model of personal informatics: the case of a quantified self. Comput Human Behav 2017 Apr;69:62-74. [doi: 10.1016/j.chb.2016.12.019]
67. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Dec 12;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
68. Kang H, Lou C. AI agency vs. human agency: understanding human–AI interactions on TikTok and their implications for user engagement. J Comput-Mediat Commun 2022 Aug 18;27(5). [doi: 10.1093/jcmc/zmac014]
69. Jung EH, Sundar SS. Senior citizens on Facebook: How do they interact and why? Comput Human Behav 2016 Aug;61:27-35. [doi: 10.1016/j.chb.2016.02.080]
70. Wilson G, Gates JR, Vijaykumar S, Morgan DJ. Understanding older adults' use of social technology and the factors influencing use. Ageing Soc 2023 Jan;43(1):222-245. [doi: 10.1017/S0144686X21000490]
71. Zhao W, Kelly RM, Rogerson MJ, Waycott J. Older adults using technology for meaningful activities during COVID-19: an analysis through the lens of self-determination theory. Presented at: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23); Apr 23-28, 2023; Hamburg, Germany p. 1-17. [doi: 10.1145/3544548.3580839]
72. Yang CC, Liu C, Wang YS. The acceptance and use of smartphones among older adults: differences in UTAUT determinants before and after training. Library Hi Tech 2023 Oct 31;41(5):1357-1375. [doi: 10.1108/LHT-12-2021-0432]
73. Chen J, Wang T, Fang Z, Wang H. Research on elderly users' intentions to accept wearable devices based on the improved UTAUT model. Front Public Health 2022;10:1035398. [doi: 10.3389/fpubh.2022.1035398] [Medline: 36699866]
74. Population in brief. National Population and Talent Division. 2024 Sep 24. URL: https://www.population.gov.sg/media-centre/publications/population-in-brief [accessed 2025-02-25]
75. Kagawa Singer M. Applying the concept of culture to reduce health disparities through health behavior research. Prev Med 2012 Nov;55(5):356-361. [doi: 10.1016/j.ypmed.2012.02.011]
76. Lee SG, Trimi S, Kim C. The impact of cultural differences on technology adoption. J World Bus 2013 Jan;48(1):20-29. [doi: 10.1016/j.jwb.2012.06.003]
77. Li Y, Rho EHR, Kobsa A. Cultural differences in the effects of contextual factors and privacy concerns on users' privacy decision on social networking sites. Behav Inf Technol 2022 Feb 17;41(3):655-677. [doi: 10.1080/0144929X.2020.1831608]

Abbreviations

AI: artificial intelligence
CFI: comparative fit index
H: hypothesis
RFID: radio frequency identification
RMSEA: root mean square error of approximation
RQ: research question
SDT: self-determination theory
SEM: structural equation model
SRMR: standardized root mean square residual
TAM: technology acceptance model
TLI: Tucker-Lewis index
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by Q Fan; submitted 19.11.24; peer-reviewed by J Farias, MA Azmi Murad; revised version received 28.01.25; accepted 30.01.25; published 19.03.25.

Please cite as:
Kang H, Yang T, Banu N, Ng SWT, Lee JK
Exploring Smart Health Wearable Adoption Among Singaporean Older Adults Based on Self-Determination Theory: Web-Based Survey Study
JMIR Aging 2025;8:e69008
URL: https://aging.jmir.org/2025/1/e69008
doi:10.2196/69008

A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study

Melissa Johanna Böttinger[1,2], MSc; Sabato Mellone[3], PhD; Jochen Klenk[4,5], Prof Dr; Carl-Philipp Jansen[1], PhD; Marios Stefanakis[6], MSc; Elena Litz[1,2], MSc; Anastasia Bredenbrock[2], BSc; Jan-Philipp Fischer[7], MA; Jürgen M Bauer[1,2], MD; Clemens Becker[1], MD; Katharina Gordt-Oesterwind[1,2,8], PhD

[1]Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Bergheimer Str. 20, Heidelberg, Germany

[2]Network Aging Research, Heidelberg University, Heidelberg, Germany

[3]Department of Electrical, Electronic and Information Engineering, University of Bologna, Bologna, Italy

[4]Institute of Epidemiology and Medical Biometry, Ulm University, Ulm, Germany

[5]Study Center Stuttgart, IB University of Health and Social Sciences, Stuttgart, Germany

[6]Department of Clinical Gerontology and Geriatric Rehabilitation, Robert Bosch Hospital, Stuttgart, Germany

[7]TSG ResearchLab gGmbH, Zuzenhausen, Germany

[8]Institute of Sports and Sports Sciences, Heidelberg University, Heidelberg, Germany

Corresponding Author:
Melissa Johanna Böttinger, MSc
Geriatric Center, Medical Faculty Heidelberg, Heidelberg University, Bergheimer Str. 20, Heidelberg, Germany

Abstract

Background: The Timed Up and Go test (TUG) is recommended as an evidence-based tool for measuring physical capacity. Instrumented TUG (iTUG) approaches expand classical supervised clinical applications offering the potential of self-assessment for older adults.

Objective: This study aimed to evaluate the concurrent validity and test-retest reliability of a smartphone-based TUG self-assessment "up&go app."

Methods: A total of 52 community-dwelling older adults (>67 years old) were recruited. A validated and medically certified system attached with a belt at the lower back was used as a reference system to validate the "up&go app" algorithm. The participants repeated the TUG 5 times wearing, a smartphone with the "up&go app" in their front trouser pocket and an inertial sensor to test the concurrent validity. A subsample of 37 participants repeated the "up&go app" measurement 2 weeks later to examine the test-retest reliability.

Results: The correlation between the "up&go app" and the reference measurement was r=0.99 for the total test duration and r=0.97 for the 5 single repetitions. Agreement between the 5 repetitions was intraclass correlation coefficient (ICC)=0.9 (0.84 - 0.94). Leaving out the first repetition, the agreement was ICC=0.95 (0.92 - 0.97). Test-retest agreement had an ICC=0.79 (0.53 - 0.9).

Conclusions: The duration of 5 repetitions of the TUG test, measured with the pocket-worn "up&go app," was very consistent with the results of a lower-back sensor system, indicating excellent concurrent validity. Participants walked slower in the first round than in the other 4 repetitions within a test run. Test-retest reliability was also excellent. The "up&go app" provides a useful smartphone-based approach to measure 5 repetitions of the TUG. The app could be used by older adults as a self-screening and monitoring tool of physical capacity at home and thereby help to early identify functional limitations and take interventions when necessary.

(JMIR Aging 2025;8:e67322) doi:10.2196/67322

KEYWORDS

timed up and go test; self-assessment; instrumented assessment; technology-based assess-ment; physical capacity; mobility; aged; mobile applications; smartphone; diagnostic self evaluation

Introduction

The baby boomer generation embraces smartphone technology [1]. This particular cohort is willing and capable of using mobile health applications especially when they are designed in a user-centered way and increase the likelihood of a self-determined future [2-4]. Good examples are the use of fitness trackers, heart rate monitoring, and glucose monitoring [5].

The World Health Organization (WHO) and United Nations (UN) member states have declared the current decade as the UN "Decade of Healthy Ageing" (2021 - 2030) to improve the lives of older people, their families, and the communities in which they live [6-8]. Early identification of health risks and functional limitations in older adults is crucial to enable timely preventive measures and promote the maintenance of independence [9]. A recent systematic review was conducted to guide WHO recommendations and to establish a "standard" for evidence-based assessment of physical capacity [10]. It recommends the Timed Up and Go Test (TUG), in addition to the 30-Second Chair Rise Test, due to the high quality evidence for its sufficiently valid and reliable measurement of physical capacity in community-dwelling older adults [10,11]. The standard TUG captures the time recorded by a trained test administrator using a stopwatch. The test person rises from a chair, walks 3 meters at their usual pace, turns around, walks back, and sits down again. Longer test durations indicate limitations in physical capacity, particularly in strength and balance [10-12] and predict hospitalization and associated functional decline in older adults [13]. The need for and lack of a digital version of the TUG has been addressed during the meetings of the WHO Locomotor Capacity Working Group. This is even more relevant for persons living remotely or in low and middle-income countries where smartphone technology is often available but health care professionals are often lacking [2].

Wearable technologies offer opportunities to lower the barrier for integrating functional tests like the TUG into the assessment and clinical management of older people [14,15]. Such an instrumented Timed Up and Go Test (iTUG) not only enables standardized and digitized measurement in supervised settings [16-18] but also opens an option for low-cost self-screening and monitoring in people's everyday lives [19,20]. In clinical studies, camera-based systems, inertial sensors or smartphones that are attached to the back (eg, with a belt) are most commonly used to perform the iTUG [21]. For self-assessment purposes, however, an approach using common technologies and a convenient placement on the body would be more feasible, without the need for additional materials or training. Therefore, the "up&go" iTUG was developed as a self-test for older adults. The prototype of the "up&go" app emerged from several EU projects [17-19,22,23] and a cocreation process with older adults [24]. It incorporates an innovative approach: placing a smartphone in the user's front trouser pocket during 5 repetitions of the test. An unsupervised smartphone-based iTUG could fill a gap in the landscape of mobility assessment methods that is still dominated by patient-reported outcome measures and supervised assessments conducted in laboratory or clinical settings [25]. Digital self-assessments could support older adults to take up an active role in risk screening or monitoring, and to early identify changes in their mobility. If necessary, this would allow them to initiate further medical diagnostics, and to take primary or secondary preventive measures in a timely manner. Thus, the motivation for this study is to examine if the "up&go" self-assessment app fulfils the criteria of validity and reliability to be considered trustworthy and to serve as a credible tool for implementation in future iTUG studies and in the self-management of older people.

Hence, this study aims to examine the concurrent validity and test-retest reliability of the "up&go" smartphone-based TUG self-assessment for older adults. The measures considered are the total TUG test duration and the duration of the 5 individual repetitions. This study corresponds to an analytical validation according to the V3 framework [26].

The objectives of this study are to analyze the concurrent validity of the algorithm used in the "up&go app" against a validated and medically certified sensor-based system and to determine the test-retest reliability of the "up&go app" algorithm in the home environment of older adults.

Methods

Participant Recruitment

The findings are reported following the guidelines for reporting reliability and agreement studies [27]. Participant were recruited from the SMART-AGE randomized controlled trial conducted at Heidelberg University, Germany [28]. Participants were included in the SMART-AGE study if they were 67 years or older, lived in the community, had basic knowledge of using PCs or tablets and of the German language. Exclusion criteria were severe medical conditions (ie, heart failure with shortness of breath at rest, cardiac arrhythmia with dizziness, Parkinson disease with use of a walker or wheelchair, cancer with chemotherapy or radiotherapy, chronic lung disease with oxygen therapy, and planned major medical procedure with inpatient hospitalization within the next 3 months), severe visual or hearing impairment, and severe cognitive impairment.

Data for the sample description were extracted from the SMART-AGE initial assessment data: age, sex, BMI, citizenship, living situation, employment, WHO Quality of Life Scale score [29], Satisfaction With Life Scale score [30], fall history, Groll Functional Comorbidity Index [31], Trail Making Test A and B score [32], Six-item Cognitive Impairment Test score [33], Short Falls Efficacy Scale International score [34], 4-Meter Walk Test score [35], 30-second Chair Rise Test score [36], and stopwatch-measured Timed-Up and Go Test score [11].

Ethical Considerations

The SMART-AGE study and the amended protocol for this sub study were approved by Heidelberg University Medical Faculty's Ethical Committee (S-672/2022). All participants provided informed signed consent before participating.

Data Collection and Processing

The data collection was performed during a home visit 3 to 4 months after the SMART-AGE study's baseline assessment. A total of 3 pretrained assessors with a professional background in psychology (AB and EL) or physiotherapy (MJB) conducted the home visits following a standardized manual.

"Up&go App": Smartphone-Based Timed Up and Go Test

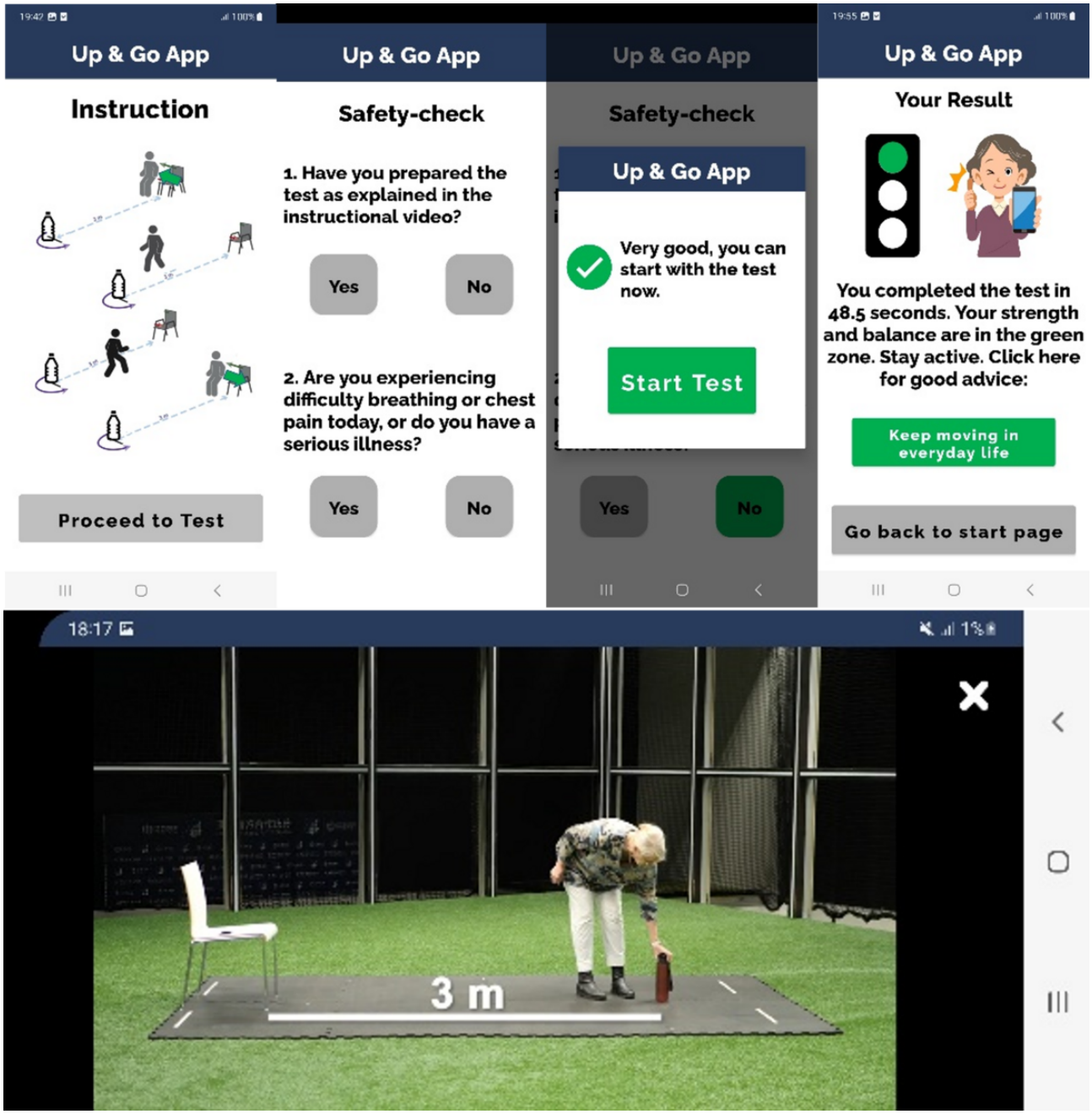
Previous versions of the "up&go app" algorithm have been developed and validated within the 2 European Commission funded projects, Farseeing [17,22] and PreventIT [18,19,23].

The “up&go app” investigated in this study [37] is a further development of the previous versions. It is based on the results of a study examining the usability of the previous Norwegian version of the PreventIT project [19] and a cocreation study with older adults [24].

In order to run the TUG test within the “up&go app,” the user is invited to watch and read the test instructions, prepare the test setting (chair, 3 m distance, marker), answer 2 safety questions, press the “start test” button, place the smartphone in one of the front trouser pockets and sit down to begin with the first round of the test at usual gait speed (Figure 1). The app includes an algorithm processing real-time accelerometer data to detect the body movements to guide the user through 5 consecutive repetitions of the TUG via audio prompts (eg, “Please put the phone in your front trouser pocket and sit down”). After each run there is a pause of about 3 seconds in which the algorithm detects the position change, and the new audio announcement is played (eg, “This was the third run. Please wait for the next start signal”). The embedded sensors in the smartphone collect data during the test.

After the test, the algorithm automatically postprocesses accelerometer data and computes the total time needed to complete 5 repetitions, excluding pauses between runs. The total time is shown on the smartphone screen in seconds together with a traffic light (Figure 1, green=<60 s, yellow 60 - 90 s, red>90 s) and an according recommendation for action (eg, green light: recommendation to stay active, yellow light: recommendation for supervised training, and red light: recommendation to seek medical advice). The color coding thresholds are based on the cut-off value of 12 seconds for 1 repetition, respectively 60 seconds for 5 repetitions, to indicate normal versus below normal mobility in community-dwelling older adults [13,38].

Figure 1. Screenshots from the “up&go app".

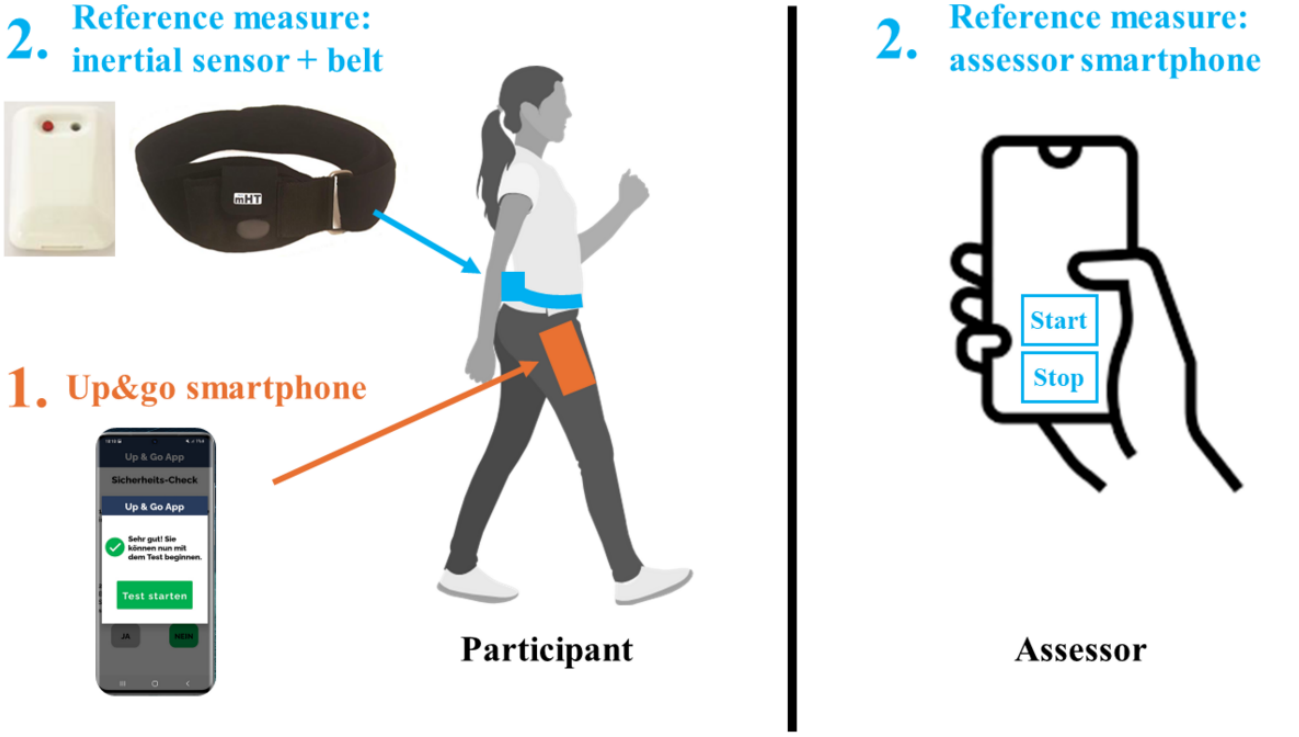


Concurrent Validity

In this study, the “mTUG” system (“mTUG” medical device, mHealth Technologies) was used as a reference measure to validate the “up&go app” algorithm. “mTUG” is a sensor-based system that has been validated in a population of older adults [17] and is certified as a medical device. The system consists of an inertial sensor (mHealth Technologies triaxial accelerometer range±2 g, triaxial gyroscope range±250°/sec, sampling rate: 100 Hz) connected by Bluetooth to an assessor smartphone (Samsung Galaxy S10e, Android 5.0.1) including a customized app (mHealth Technologies) allowing the assessor to manually start and stop the sensor measurement.

The TUG was repeated 5 times with participants wearing the 2 measurement systems simultaneously (Figure 2), smartphone with “up&go app” in their right front trouser pocket and an inertial sensor (mHealth Technologies) attached to an elastic belt worn on the lower back. A previous study with hospitalized hip fracture patients found that the performance of three TUG trials is recommended to achieve performance stability [39]. In this study, 5 repetitions were performed to take into account more variability in the performance of older adults conducting the test at home.

Figure 2. : Simultaneous measurement with (1) the "up&go app" and (2) the reference measure.



Participants were familiar with the TUG procedure as the conventional stopwatch TUG [11] had been performed beforehand as part of the baseline motor assessments within the SMART-AGE study. As the feasibility of the "up&go app" self-assessment was not part of this study, test preparation, viewing of instructional videos by participants were omitted for data collection. The test procedure was explained to the study partner (5 repetitions, usual walking pace, and follow audio instructions). The assessor prepared the TUG test setting, mounted the sensor in the belt on the participant's lower back and started the test on the "mTUG" assessor smartphone. Then, the assessor started the test on the "up&go" smartphone and handed it over to the participant. The participant then placed it in the right front trouser pocket and performed the 5 repetitions of the TUG test, following the audio instructions of the "up&go app."

The "up&go app" does not store any data on the user's smartphone to guarantee data confidentiality. We therefore developed an adapted version for this study, allowing the app to store all raw data collected from the accelerometer and gyroscope embedded in a Samsung Galaxy S21 smartphone (Android 5.0.1, triaxial accelerometer range±8 g, triaxial gyroscope range±1000°/sec, sampling rate 500 Hz) on the smartphone memory. This allowed postprocessing of the raw data and analysis of not only the total test duration, but also the duration of each of the 5 repetitions (rep1, rep2, rep3, rep4, and rep5).

Within the "mTUG" system, raw data was automatically stored on the assessor smartphone's memory and processed by a validated algorithm [17].

Test-Retest Reliability

To investigate the test-retest reliability of the "up&go app" algorithm, we repeated the "up&go" measurement at an additional home visit within 2 weeks after the first measurement, following the same procedure as described above.

Participants were asked to participate in the retest if they lived <20 km from the study center to reduce the burden of assessor time.

Statistical Analysis

The sample size for the validation part of the study was calculated using R (version 2023.09.1, R Foundation for Statistical Computing, package "pwr"). A strong correlation of ≥0.75 was assumed. With a power of 95% and a statistical significance level of α=.05, ≥45 participants are required. The sample size for the test-retest part of the study was calculated using R (version 2023.09.1, package "ICC.sample.size"). An intraclass correlation coefficient (ICC) of ≥0.75 was assumed. With power=95% and α=.05, ≥37 participants were required for 2 measurements (first measurement vs retest). Assuming a 15% dropout rate, 52 participants were recruited for the study. From this sample, we consecutively recruited participants for the retest measurement until a subsample size of at least 37 participants was reached to examine the test-retest reliability of the "up&go app" algorithm. The distribution of the data was checked before analyses (Shapiro-Wilk test). All data fulfilled the normal distribution assumption. For all statistical tests, α=.05 was used as the threshold for statistical significance. Pearson correlation coefficients were calculated to examine the association between the systems ("up&go app" and reference measure) regarding the total duration needed to complete the test as well as the duration of the 5 single repetitions. Correlation coefficients of 0 - 0.19 are interpreted as very weak, 0.2 - 0.39

as weak, 0.4 - 0.59 as moderate, 0.6 - 0.79 as strong, and 0.8 - 1 as very strong [40]. For graphical description of the agreement between the 2 systems ("up&go app" and reference measure), Bland-Altman plots were used [41], including the lower and upper limits of agreement (LLoA and ULoA). From the data collected during the first measurement ("up&go app" and reference measure), ICC 3,1 were calculated to analyze the agreement between the 5 repetitions. Test-retest reliability of the "up&go app" was evaluated using ICC [1,2,42]. An ICC value<0.4 indicates poor reliability, an ICC value between ≥0.4 and <0.75 fair to good reliability and an ICC value ≥0.75 indicates excellent reliability [43]. It was assumed that the association between "up&go app" and the reference measure is very strong (r>0.8) and the agreement between "up&go" and the reference measure is good (Bland-Altman plot). Furthermore, we assumed that the test-retest reliability of the "up&go" test meets excellent levels (ICC≥0.75). Statistical analyses were computed using statistical software R (version 2023.09.1) and MATLAB (version R2022b, MathWorks).

Results

Participant Data

A total of 52 community-dwelling older adults aged between 66 and 88 years (mean 73.6, SD 5.4) were measured, with a similar distribution of men (25/52, 48%) and women (27/52, 52%) and a mean BMI of 27.1 kg/m^2 (SD 4.1). Of those who provided the data, all were retired, almost all were German citizens (49/52, 98%) and one third (16/52, 33%) lived alone. Participants indicated an average of 2.3 (SD 1.7) comorbidities according to the Groll Index. Of these, degenerative disc disease (17/52), visual impairment (13/52), depression (11/52), and osteoarthritis (11/52) were reported most frequently. Cognition was normal (<7 errors on the Six-item Cognitive Impairment Test) and executive functions (Trail Making Test A and B) were in accordance with normative data of this age group [44]. The average duration of the TUG measured with the stopwatch in the SMART-AGE initial assessment was 9.7 (SD 1.7) seconds. None of the study participants used a walking aid during the TUG. Table 1 provides a detailed description of the sample.

Table . Characteristics of the study sample.

Data from the SMART-AGE initial assessment[a]		Values	Available sample
Age (years), mean (SD)		73.6 (5.4)	52
Female sex, n (%)		27 (52)	52
BMI (kg/m^2), mean (SD)		27.1 (4.1)	52
German citizenship, n (%)		49 (98)	50
Living alone, n (%)		16 (33)	49
Retired, n (%)		50 (100)	50
Stopwatch-measured TUG (s), mean (SD)		9.7 (1.7)	52
4-meter walk test, mean (SD)			
	Duration (s)	4.2 (0.9)	52
	Gait speed (m/s)	1.0 (0.2)	52
30-Second chair rise test, mean (SD)			
	Number of repetitions	12.3 (3.3)	52
Short Falls Efficacy Scale International, mean (SD)			
	Score	8.2 (2.2)	49
Groll Functional Comorbidity Index (score), mean (SD)		2.3 (1.7)	48
Fall history ("Yes, I fell in the last 12 months")[b], n (%)		19 (37)	51
Trail making test, mean (SD)			
	Test duration A (s),	45.2 (16.6)	52
	Test duration B (s)	99 (50.3)	52
	Ratio (B/A)	2.3 (1.1)	52
Six-item cognitive impairment test [c] , mean (SD)			
	Score	1.0 (1.5)	52
Satisfaction with life scale, mean (SD)			
	Score	5.0 (1.2)	49
WHO Quality of Life Scale, mean (SD)			
	Physical domain (score)	77.3 (13.1)	49
	Psychological domain (score)	72.4 (10.3)	49
	Social relationships domain (score)	65.3 (13.6)	49
	Environment domain (score)	79.8 (10.1)	49

[a]The initial SMART-AGE assessment took place 3-4 months before the data collection for this study.

[b]Fall history data was obtained one month after the initial SMART-AGE assessment.

[c]Participants with >7 error points were excluded from the SMART-AGE study due to a suspected dementia disorder.

Concurrent Validity

In total, 35 (67.3%) participants completed the test successfully on the first attempt. Due to technical problems (eg, smartphone shifted in trouser pocket), 12 (23.1%) participants had to repeat the test a second time, 4 (7.7%) participants a third time, and 1 (1.9%) participant a fourth time. After an unsuccessful attempt, participants always took a break of at least 5 minutes. As the app only saves data from complete test runs, failed attempts were not included.

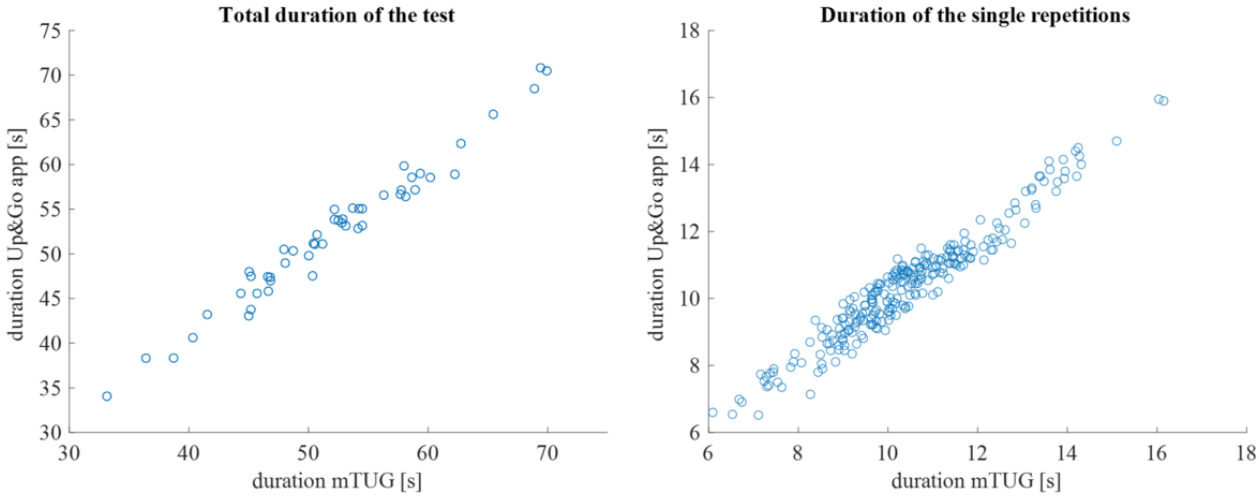
On average, participants needed 52.4 (SD 7.8) seconds to complete 5 repetitions of the TUG, as measured by the "up&go app". Table 2 shows the results for the total duration of the test and the duration of the 5 repetitions as measured by the "up&go app" and the reference measure.

Data from all 52 participants were available for the concurrent validity analysis ("up&go app" vs reference measure). The Pearson correlation coefficient (r) for the total duration was 0.99 (95% CI 0.98 - 0.99) and that for the durations of the 5 single repetitions was 0.97 (95% CI 0.96-.97; Figure 3).

Table . Total test duration and duration of the 5 repetitions measured by the "up&go app" and reference measure.

Duration	"up&go app", mean (SD)	Reference measure, mean (SD)
Total duration of the test (s)	52.38 (7.8)	52.31 (8.11)
First repetition (s)	10.99 (2.1)	11.11 (2.17)
Second repetition (s)	10.33 (1.54)	10.39 (1.61)
Third repetition (s)	10.25 (1.57)	10.34 (1.64)
Fourth repetition (s)	10.21 (1.52)	10.2 (1.52)
Fifth repetition (s)	10.11 (1.46)	10.26 (1.6)

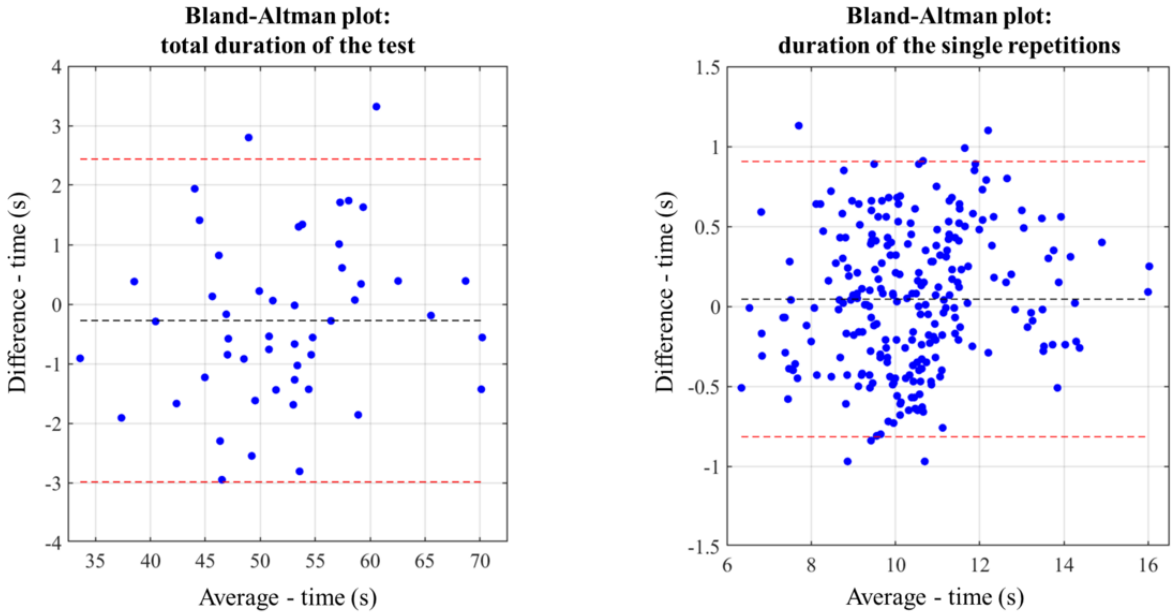
Figure 3. Correlation analysis between "up&go app" and reference measure: total duration and duration of the single repetitions. TUG: Timed Up and Go Test.



The agreement between the total duration measured by the "up&go app" and the reference measure is shown in Figure 4 (Bland-Altman plots). According to Bland-Altman analysis, the mean difference was −0.27 for the total duration of the test and −0.05 for the duration of the single repetitions. The difference in the total duration of the test was between −2.99 seconds (LLoA) and −2.44 seconds (ULoA) and the difference in the duration of a single repetition was within −0.82 seconds (LLoA) and −0.91 seconds (ULoA) with respect to the estimate of the reference measure.

Agreement between all "up&go app" repetitions was ICC(3,1)=0.9 (95% CI 0.84 - 0.94) and ICC(3,1)=0.87 (95% CI 0.80 - 0.93) between all repetitions measured by the reference measure. Leaving out the first repetition, the agreement of repetitions 2 - 5 was ICC(3,1)=0.95 (95% CI 0.92 - 0.97) for the "up&go app" and ICC(3,1)=0.93 (95% CI 0.89 - 0.96) for the reference measure.

Figure 4. Statistical agreement between the "up&go app" and the reference measure—Bland-Altman plot.



Test-Retest Reliability

Of the 52 participants, 37 were measured twice, and 1 participant was excluded due to noncompliance with the retest instruction to walk at habitual gait speed. Retests were performed on average 6 (SD 2.6) days after the first measurement. At the "up&go app" retest measurement, participants needed on average 48.6 (SD 8.5) seconds to complete the 5 repetitions. The ICC between the total duration during the first home visit and during the retest was ICC(2,1)=0.79 (95% CI 0.53 - 0.9).

Discussion

Principal Findings

To enable self-assessment for older people in their own homes, the "up&go app" was developed as a pocket-worn approach with 5 repetitions of the TUG. This study aimed to investigate the concurrent validity and test-retest reliability of the "up&go app," which records and processes data from sensors embedded in a smartphone that is placed in the front pocket of the trousers. The results show excellent agreement between the "up&go" data and the comparator system, a previously validated medically certified lower-back inertial sensor system. Overall, the correlation of the 5 single repetitions was very strong. The results indicate that participants did not seem to walk at the same speed in each of the 5 test rounds, that is, they tended to walk slower during the first round. This trend probably shows a learning or accustoming effect. Similar observations have been made in previous studies [39,45] where participants started rather slowly and more carefully during the first attempt of motor performance tests to follow the instructions precisely and to avoid errors [45]. This first round may therefore be comparable to a "trial run," which is often performed in assessments before the actual measurement. Performing several repetitions of the test provides the opportunity to observe any changes in walking speed during the test and, if necessary, to interpret it clinically regarding physical capacity. For example, in more frail populations, a significant reduction in walking speed during the test could indicate fatigue. Furthermore, this approach might achieve a better approximation toward the "real" usual walking speed compared with measuring only 1 or 2 rounds. Using the app to monitor differences in the total test duration across several measurements would be a clinically relevant use case. But what would be considered a "relevant change"? Minimal important differences (MID) of >2 seconds for 1 TUG repetition were reported in studies observing populations with age-associated disorders and TUG baseline values around 20 seconds [46,47]. In another study observing older adults with hip osteoarthritis and a TUG baseline duration of 7 seconds, a sample that is comparable with this study, the MID is significantly lower with, 0.8 seconds [48]. This suggests that ceiling effects make it more difficult to screen for relevant changes in more robust individuals. The "up&go app," however, should not only aim to screen for subjectively noticeable changes, but also for changes falling below the abovementioned MID thresholds. Looking at the Bland-Altman plots, a clinically relevant change could be assumed if the measured change is outside the limits of agreement. The limits identified in this study were about ±3 seconds (LLoA, ULoA) for the total duration of 5 repetitions and about ±1 second (LLoA, ULoA) for a single round. Especially in relatively fit, prefrail older adults, a 5-repetition approach could be more sensitive to capture small changes and clinically more useful for early identification of declining physical capacity.

The test-retest reliability analysis showed excellent results. Looking at the average total test duration, we observed a shorter total test duration in the retest. On average, retests took place 6 days after the initial measurement, which makes an improved TUG test performance through training effects unlikely. A similar observation was described elsewhere [45]. A learning effect could account for the shorter duration of the retest. Since participants were already familiar with the test from the first measurement, they may have been more confident and faster during the retest. Other iTUG systems measuring at the height of the navel, (ICC=0.97) [49] and the lower back (ICC=0.9 - 0.96) [50] show slightly higher test-retest reliability

values. This is to be expected, as the “up&go app” pocket position, specifically selected for the self-assessment purpose, is located more distant from the center of mass. Another reason could be that individuals wore different clothes with different types of pockets; these pockets may be tight or loose, which aggravates individual differences.

For future monitoring purposes, machine learning could be implemented to enable an in-app plausibility check. This might be a helpful feature to detect implausible changes in test duration (eg, >20% compared with previous test performances very recently).

Limitations

The sample was predominantly White, and participants were relatively fit, as demonstrated by the TUG performance. The average time needed to complete one repetition (9.7, SD 1.7 seconds; Table 1) is below the threshold of 12 seconds to distinguish normal versus below normal mobility [38]. The results are therefore limited in their transferability to more diverse cultural populations with greater mobility restrictions. For pragmatic reasons, we did not include a gold-standard reference measure, that is, optoelectronic measurement, limiting the results of this study accordingly. A segmental analysis of the “up&go app” data (sit-to-stand, walking, turning, and stand-to-sit) was not considered to be useful for self-monitoring. Furthermore, signal noise caused by the pocket-worn approach would be expected, as the smartphone moves back and forth in the trouser pocket in addition to the body movement during the test. Thus, the pocket-worn approach should be considered as a consumer-centered approach and not as a measurement system for segmental or kinematic analysis. For an accurate segmentation of the test, the use of technology that is placed close to the center of mass would be recommended. The instruction to walk at “usual walking speed” leaves room for individual interpretation, hampering standardization. The extent to which the gait speed during the TUG test corresponds to an individual’s actual real-world gait speed was not assessed in this study. It is assumed that performing 5 repetitions and the familiar home setting enabled a greater approximation to the normal walking speed. However, an activity measurement of several days would be required for comparison [51].

Future Perspectives

The smartphone app used in this study could be suitable as a physical capacity screening and monitoring tool incorporating an instrumented approach for the widely used TUG test. It is designed to be used by older adults, but it could also be implemented by health care professionals to measure TUG total time in clinical settings.

As this study was aimed at validating the test algorithm, it is currently not possible to draw conclusions on the feasibility and user experience of older adults when operating the app, setting up and conducting the test independently. The app should therefore be examined and further developed in future cocreation studies with the target group.

The investigated app is designed to screen physical capacity, representing one of several health domains. In the future, it could be used as part of a comprehensive digital self-assessment, which should include other relevant risk factors such as physical activity, cognition, and vision in addition to physical capacity [24,52]. Self-guided early detection of risks in these domains could enable timely, more specific clinical diagnostics and the initiation of appropriate care interventions. Implementing the app as an upstream assessment within a digital training platform would allow adopting the dosage and supporting the selection of target-specific exercises to be tailored to the user’s individual level of physical capacity.

Conclusion

The results show excellent concurrent validity and test-retest reliability of the pocket-worn iTUG approach with 5 repetitions. The “up&go app” could be suitable as a self-screening and self-monitoring of physical capacity for older adults at home. It provides a smartphone-based approach to accurately measure the total duration of TUG. This novel approach offers the potential for older adults to take an active role in their health management and preclinical risk detection.

Acknowledgments

This publication is based on data from the study "Smart Aging in Community Contexts: Testing Intelligent Assistive Systems for Self-regulation and Co-regulation under Real-Life Conditions" (SMART-AGE). The study is funded by the Carl-Zeiss-Foundation. We thank all project members from the SMART-AGE project and the TSG Research Lab for their contribution. In particular, the authors extend thanks to Carlo Tacconi, Jan Spielmann, Jan Mayer, Ronny Bergquist, Sarah Zimmermann, Tobias Eckert, Claudia Hellmund, Franziska Kramer-Gmeiner, Meike Steinhilber, Anna-Lena Schubert and Alica Mertens.

Authors' Contributions

MJB, CB, and KGO conceptualized the study. MJB, AB, EL, and MS took part in data collection. MJB and SM took part in the data analysis. MJB drafted the manuscript, JK, SM, CB, and KGO took part in a preliminary process of reviewing and editing the manuscript. All authors participated in the subsequent editing process, read the entire manuscript, and agreed to the published version.

Conflicts of Interest

SM holds a share of the mHealth Technologies srl company. None of the other authors declared conflicts of interest. The “up&go app” is freely available in the Google Play and Apple App Store, in German speaking countries.

References

1. Pew Research Center. Share of those 65 and older who are tech users has grown in the past decade. 2022. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-10-09]
2. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8] [Medline: 37464029]
3. Fotteler M, Risch B, Gaugisch P, et al. Obstacles to using assistive technology for older adults - results from a focus group analysis. Stud Health Technol Inform 2021 May 27;281:994-998. [doi: 10.3233/SHTI210327] [Medline: 34042822]
4. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
5. Malani P, Kullgren J, Solway E, et al. National Poll on Healthy Aging: Mobile Health App Use Among Older Adults. 2022.
6. United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2019: highlights. [(ST/ESA/SER.A/430)]. 2019.
7. World Health Organization. Decade of healthy ageing: plan of action. 2020. URL: https://www.who.int/publications/m/item/decade-of-healthy-ageing-plan-of-action [accessed 2024-10-09]
8. Amuthavalli Thiyagarajan J, Mikton C, Harwood RH, et al. The UN Decade of healthy ageing: strengthening measurement for monitoring health and wellbeing of older people. Age Ageing 2022 Jul 1;51(7):afac147. [doi: 10.1093/ageing/afac147] [Medline: 35776669]
9. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4]
10. Honvo G, Sabico S, Veronese N, et al. Measures of attributes of locomotor capacity in older people: a systematic literature review following the COSMIN methodology. Age Ageing 2023 Oct 28;52(Suppl 4):iv44-iv66. [doi: 10.1093/ageing/afad139] [Medline: 37902521]
11. Podsiadlo D, Richardson S. The timed “Up & Go”: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
12. Coelho-Junior HJ, Rodrigues B, Gonçalves IDO, Asano RY, Uchida MC, Marzetti E. The physical capabilities underlying timed “Up and Go” test are time-dependent in community-dwelling older women. Exp Gerontol 2018 Apr;104:138-146. [doi: 10.1016/j.exger.2018.01.025] [Medline: 29410234]
13. Gatenio-Hefling O, Tzemah-Shahar R, Asraf K, Dilian O, Gil E, Agmon M. Revisiting the “Timed Up and Go” test: a 12-s cut-off can predict Hospitalization Associated Functional Decline in older adults. Geroscience 2024 Jul 16. [doi: 10.1007/s11357-024-01280-3] [Medline: 39014130]
14. Ponciano V, Pires IM, Ribeiro FR, et al. Is the Timed-Up and Go Test feasible in mobile devices? A systematic review. Electronics (Basel) 2020;9(3):528. [doi: 10.3390/electronics9030528]
15. Böttinger MJ, Labudek S, Schoene D, et al. “TiC-TUG”: technology in clinical practice using the instrumented timed up and go test-a scoping review. Aging Clin Exp Res 2024 Apr 27;36(1):100. [doi: 10.1007/s40520-024-02733-7] [Medline: 38676844]
16. Milosevic M, Jovanov E, Milenkovic A. Quantifying timed-up-and-go test: A smartphone implementation. Presented at: 2013 IEEE International Conference on Body Sensor Networks (BSN 2013): Cambridge, MA, USA; May 6-9, 2013. [doi: 10.1109/BSN.2013.6575478]
17. Mellone S, Tacconi C, Chiari L. Validity of a Smartphone-based instrumented Timed Up and Go. Gait Posture 2012 May;36(1):163-165. [doi: 10.1016/j.gaitpost.2012.02.006] [Medline: 22421189]
18. Bergquist R, Nerz C, Taraldsen K, et al. Predicting advanced balance ability and mobility with an instrumented Timed Up and Go Test. Sensors (Basel) 2020 Sep 3;20(17):4987. [doi: 10.3390/s20174987] [Medline: 32899143]
19. Bergquist R, Vereijken B, Mellone S, Corzani M, Helbostad JL, Taraldsen K. App-based self-administrable clinical tests of physical function: development and usability study. JMIR Mhealth Uhealth 2020 Apr 27;8(4):e16507. [doi: 10.2196/16507] [Medline: 32338616]
20. Fudickar S, Pauls A, Lau S, et al. Measurement system for unsupervised standardized assessments of Timed Up and Go Test and 5 Times Chair Rise Test in community settings-a usability study. Sensors (Basel) 2022 Jan 19;22(3):731. [doi: 10.3390/s22030731] [Medline: 35161478]
21. Ortega-Bastidas P, Gómez B, Aqueveque P, Luarte-Martínez S, Cano-de-la-Cuerda R. Instrumented Timed Up and Go Test (iTUG)-more than assessing time to predict falls: a systematic review. Sensors (Basel) 2023 Mar 24;23(7):3426. [doi: 10.3390/s23073426] [Medline: 37050485]
22. Mellone S, Tacconi C, Schwickert L, Klenk J, Becker C, Chiari L. Smartphone-based solutions for fall detection and prevention: the FARSEEING approach. Z Gerontol Geriat 2012 Dec;45(8):722-727. [doi: 10.1007/s00391-012-0404-5]

23. Coni A, Ancum JMV, Bergquist R, et al. Comparison of standard clinical and instrumented physical performance tests in discriminating functional status of high-functioning people aged 61 70 years old. Sensors (Basel) 2019 Jan 22;19(3):449. [doi: 10.3390/s19030449] [Medline: 30678268]
24. Böttinger MJ, Litz E, Gordt-Oesterwind K, et al. Co-creating a digital life-integrated self-assessment for older adults: user experience study. JMIR Aging 2023 Sep 26;6:e46738. [doi: 10.2196/46738] [Medline: 37751274]
25. Soubra R, Chkeir A, Novella JL. A systematic review of thirty-one assessment tests to evaluate mobility in older adults. Biomed Res Int 2019;2019:1354362. [doi: 10.1155/2019/1354362] [Medline: 31321227]
26. Goldsack JC, Coravos A, Bakker JP, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3:55. [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
27. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]
28. Federal Institute for Drugs and Medical Devices. German clinical trials register: SMART aging in community contexts: testing intelligent assistive systems for self-regulation and co-regulation under real-life conditions (SMART-AGE). 2024. URL: https://drks.de/search/en/trial/DRKS00034316 [accessed 2024-12-13]
29. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
30. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
31. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol 2005 Jun;58(6):595-602. [doi: 10.1016/j.jclinepi.2004.10.018] [Medline: 15878473]
32. Reitan RM. Trail Making Test: Reitan Neuropsychology Laboratory; 1992.
33. Gale TM, Larner AJ. Six-item cognitive impairment test (6CIT). In: Larner AJ, editor. Cognitive Screening Instruments: Springer; 2017:241-253. [doi: 10.1007/978-3-319-44775-9_11]
34. Kempen GIJM, Yardley L, van Haastregt JCM, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008 Jan;37(1):45-50. [doi: 10.1093/ageing/afm157] [Medline: 18032400]
35. Kim HJ, Park I, Lee HJ, Lee O. The reliability and validity of gait speed with different walking pace and distances against general health, physical function, and chronic disease in aged adults. J Exerc Nutrition Biochem 2016 Sep;20(3):46-50. [doi: 10.20463/jenb.2016.09.20.3.7] [Medline: 27757387]
36. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as A measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
37. Google Ireland Limited. Up&go (google play). 2024. URL: https://play.google.com/store/apps/details?id=mht.tsg.upandgo&pli=1 [accessed 2024-10-09]
38. Bischoff HA, Stähelin HB, Monsch AU, et al. Identifying a cut-off point for normal mobility: a comparison of the timed "up and go" test in community-dwelling and institutionalised elderly women. Age Ageing 2003 May;32(3):315-320. [doi: 10.1093/ageing/32.3.315] [Medline: 12720619]
39. Kristensen MT, Ekdahl C, Kehlet H, Bandholm T. How many trials are needed to achieve performance stability of the Timed Up & Go test in patients with hip fracture? Arch Phys Med Rehabil 2010 Jun;91(6):885-889. [doi: 10.1016/j.apmr.2010.01.021] [Medline: 20510979]
40. Evans JD. Straightforward Statistics for the Behavioral Sciences: Brooks/Cole Pub. Co; 1996.
41. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
42. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
43. Rosner B. Fundamentals of Biostatistics, 8th edition: Cengage Learning; 2016.
44. Tombaugh TN. Trail Making Test A and B: normative data stratified by age and education. Arch Clin Neuropsychol 2004 Mar;19(2):203-214. [doi: 10.1016/S0887-6177(03)00039-8] [Medline: 15010086]
45. Vartiainen MV, Rinne MB, Lehto TM, Pasanen ME, Sarajuuri JM, Alaranta HT. The test–retest reliability of motor performance measures after traumatic brain injury. Adv Physiother 2006 Jan;8(2):50-59. [doi: 10.1080/14038190600700195]
46. Gallagher R, Marquez J, Osmotherly P. Clinimetric properties and minimal clinically important differences for a battery of gait, balance, and cognitive examinations for the Tap Test in idiopathic mormal pressure hydrocephalus. Neurosurgery 2019 Jun 1;84(6):E378-E384. [doi: 10.1093/neuros/nyy286] [Medline: 30010977]
47. Braun T, Thiel C, Schulz RJ, Grüneberg C. Responsiveness and interpretability of commonly used outcome assessments of mobility capacity in older hospital patients with cognitive spectrum disorders. Health Qual Life Outcomes 2021 Mar 1;19(1):68. [doi: 10.1186/s12955-021-01690-3] [Medline: 33648508]

48. Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. J Orthop Sports Phys Ther 2011 May;41(5):319-327. [doi: 10.2519/jospt.2011.3515] [Medline: 21335930]
49. Yamada S, Aoyagi Y, Yamamoto K, Ishikawa M. Quantitative evaluation of gait disturbance on an instrumented Timed Up-and-go Test. Aging Dis 2019 Feb;10(1):23-36. [doi: 10.14336/AD.2018.0426] [Medline: 30705765]
50. Miller Koop M, Ozinga SJ, Rosenfeldt AB, Alberts JL. Quantifying turning behavior and gait in Parkinson's disease using mobile technology. IBRO Rep 2018 Dec;5:10-16. [doi: 10.1016/j.ibror.2018.06.002] [Medline: 30135951]
51. Warmerdam E, Hausdorff JM, Atrsaei A, et al. Long-term unsupervised mobility assessment in movement disorders. Lancet Neurol 2020 May;19(5):462-470. [doi: 10.1016/S1474-4422(19)30397-7] [Medline: 32059811]
52. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr 2014 Feb 1;14:14. [doi: 10.1186/1471-2318-14-14] [Medline: 24484314]

Abbreviations

ICC: intraclass correlation coefficient
iTUG: instrumented Timed Up and Go Test
LLoA: lower level of agreement
MID: minimal important difference
TUG: Timed Up and Go Test
ULoA: upper level of agreement
UN: United Nations
WHO: World Health Organization

Edited by AF Abdulai; submitted 18.10.24; peer-reviewed by H Kocaman, S Rego; revised version received 13.12.24; accepted 02.01.25; published 21.03.25.

Please cite as:
Böttinger MJ, Mellone S, Klenk J, Jansen CP, Stefanakis M, Litz E, Bredenbrock A, Fischer JP, Bauer JM, Becker C, Gordt-Oesterwind K
A Smartphone-Based Timed Up and Go Test Self-Assessment for Older Adults: Validity and Reliability Study
JMIR Aging 2025;8:e67322
URL: https://aging.jmir.org/2025/1/e67322
doi:10.2196/67322

Perceptions of the Use of Mobile Apps to Assess Sleep-Dependent Memory in Older Adults With Subjective and Objective Cognitive Impairment: Focus Group Approach

Aaron Lam[1,2,3], PhD; Simone Simonetti[1,2], PhD; Angela D'Rozario[3], PhD; David Ireland[4], PhD; DanaKai Bradford[4], PhD; Jurgen Fripp[4], PhD; Sharon L Naismith[1,2], PhD

[1]Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 96 Mallett Street, Camperdown, Australia

[2]School of Psychology, The University of Sydney, Camperdown, Australia

[3]Woolcock Institute of Medical Research, Macquarie University, Macquarie Park, Australia

[4]Australian eHealth Research Centre, Commonwealth Scientific and Industrial Research Organisation, Herston, Australia

Corresponding Author:
Aaron Lam, PhD
Healthy Brain Ageing Program, The Brain and Mind Centre, The University of Sydney, 96 Mallett Street, Camperdown, Australia

Abstract

Background: Sleep-dependent memory (SDM) is the phenomenon where newly obtained memory traces are consolidated from short-term memory stores to long-term memory, underpinning memory for daily life. Administering SDM tasks presents considerable challenges, particularly for older adults with memory concerns, due to the need for sleep laboratories and research staff being present to administer the task. In response, we have developed a prototype mobile app aimed at automating the data collection process.

Objective: This study investigates the perspectives of older adults, with subjective or objective cognitive impairment, regarding barriers and facilitators to using a new mobile app for at-home assessment of SDM.

Methods: In total, 11 participants aged 50 years and older were recruited from the Healthy Brain Ageing memory clinic, a specialized research memory clinic that focuses on the assessment and early intervention of cognitive decline. Two focus groups were conducted and thematically analyzed using NVivo (version 13; Lumivero).

Results: On average, participants were aged 68.5 (SD 5.1) years, and 4/11 were male. Eight participants had subjective cognitive impairment, and 3 participants had mild cognitive (objective) impairment. Two main themes emerged from the focus groups, shedding light on participants' use of mobile phones and the challenges and facilitators associated with transitioning from traditional laboratory-based assessments to home assessments. These challenges include maintaining accurate data, engaging with humans versus robots, and ensuring accessibility and task compliance. Additionally, potential solutions to these challenges were identified.

Conclusions: Our findings underscore the importance of app flexibility in accommodating diverse user needs and preferences as well as in overcoming barriers. While some individuals required high-level assistance, others expressed the ability to navigate the app independently or with minimal support. In conclusion, older adults provided valuable insights into the app modifications, user needs, and accessibility requirements enabling home-based SDM assessment.

(JMIR Aging 2025;8:e68147) doi:10.2196/68147

KEYWORDS

aging; mild cognitive impairment; subjective cognitive impairment; digital health; cognition; neuropsychology; sleep

Introduction

There are over 55 million people living with dementia [1]. Dementia is an umbrella term for a diverse number of substantial cognitive impairments that significantly interfere with an individual's daily living. The most common form of dementia is Alzheimer disease, in which memory impairment is a critical early feature [2] that is intricately linked to diminished well-being and quality of life [3]. Some facets of memory can be assessed by standardized neuropsychological tests that examine a person's ability to recall information presented to them after a short delay (eg, 20 - 30 minutes). An aspect of memory that cannot be assessed during a standardized neuropsychological test is sleep-dependent memory (SDM). This is a process by which memories encoded during the day are consolidated and strengthened during sleep [4]. This mechanism greatly influences overall memory capabilities by bolstering long-term memory storage, integration, and retrieval.

To date, most studies have explored SDM in younger adults [5]. Thus, tasks are tailored to suit the cognitive capacities of

younger individuals and often involve memorizing many items. This poses significant challenges for older adults, particularly those with cognitive impairments, leading to a notable research gap in this demographic. Given this gap and the importance of SDM in memory for daily life, our team developed an SDM task [6] tailored explicitly for older adults with cognitive concerns and mild cognitive impairment (MCI). MCI refers to a transitional state between normal aging and dementia, where cognitive impairment is apparent, but daily functioning is largely intact. Our prior work demonstrated that individuals with multiple-domain MCI had significantly more compromised SDM than healthy controls and those with single-domain MCI [6]. In this study, poorer SDM was linked to having greater sleep apnea severity for older adults without MCI. In contrast, for those with MCI, poorer performance was associated with decreased sleep spindle duration and smaller hippocampal subfield size. As various age-related sleep changes occur both naturally [7] and with neurodegenerative diseases [8,9], it is crucial to understand how these alterations impact SDM. Notably, preliminary evidence suggests that SDM may be improved through sleep apnea management [10], transcranial electrical stimulation [11], and acoustic stimulation [12]. These findings underscore the importance of identifying clinical correlates and predictors of poor SDM performance to inform targeted interventions.

Unfortunately, our understanding of SDM is limited to small sample case-control studies, and it is not routinely assessed in clinical evaluations. This may be partly attributable to the type of task used and the setting within which most research has been conducted. In addition to being predominantly in younger healthy samples, studies to date have largely been conducted in sleep laboratories where participants are asked to do memory testing before and after sleep [13]. This requires participants to sleep in an unfamiliar environment, meaning their sleep that night may not reflect their usual sleep [14], and they are often asked to learn significant amounts of information before sleep using tasks that are not feasible or suitable for clinical settings. Another barrier is the high demand for staff to conduct these studies (Naismith, personal communication, 2024). Staff must be present to administer SDM tasks before sleep (often within 3 hours before sleep) and after sleep (often after 1-hour postwaking) and to stay overnight to monitor the participant. Together, the requirement for sleep laboratories and high staff demand result in high costs associated with conducting SDM tasks. These barriers could be addressed by adapting SDM tasks for older clinical samples and delivering the tasks in the home environment.

The adoption of digital health technology in older adults has been increasing [15], including by individuals with subjective concerns about their cognition and those with MCI [16]. However, when considering new app-based approaches, barriers and facilitators that influence adoption and effective use must be identified [17]. A prior systematic review showed that barriers preventing the use of digital technology in older adults (including those with cognitive impairment) can include motor, sensory, cognitive, lack of familiarity, and device-specific challenges, though these factors can be mitigated through design modifications [18]. In contrast, digital health technology adoption can be facilitated by the perceived usefulness of the app, pre-existing knowledge, and ease of use. For SDM tasks, additional considerations are required, given that they are optimally tested under strict timings (ie, before and after sleep) and may require participants' sustained attention and interaction with a mobile app for extended periods (eg, potentially up to an hour).

This study sought to explore the perspectives of older adults with subjective or objective cognitive impairment regarding barriers to and facilitators of an at-home assessment of SDM using a new "chatbot" (mobile app), approach to assessment, named "Sleep Memories." A chatbot is designed to simulate human-like conversation with users through text or voice interactions and, in our scenario, to deliver a memory task and record a participant's responses. Previous work has shown the use of chatbots in various health populations, including chronic pain populations to deliver pain education [19], and monitoring of chemotherapy outcomes in older adults with cancer [20]. The objective was to use co-design methodologies to assess an initial app prototype, which could be incorporated into a revised version suitable for clinical testing. We were specifically interested in potential barriers to and facilitators of use.

Methods

Digitalization of the SDM Task

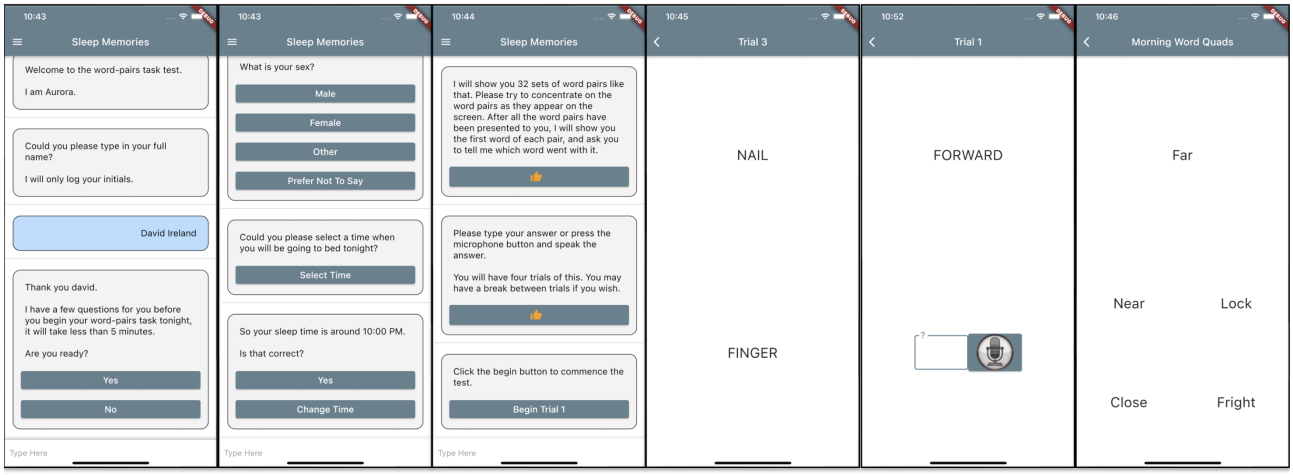
The SDM task, as described in Lam et al [6], was initially adapted to be suited for an app-based platform (Figure 1). This task is a verbal memory task that comprises 32 word pairs, half of which are semantically related and the other half unrelated. In step 1, participants are presented with all 32 pairs. In step 2, individual words are presented, for which participants are required to identify the corresponding pair word by speech or text. As described in the initial study [6], participants performed this task 4 times in the evening within 3 hours of bedtime to facilitate learning. In the morning, participants completed the recall component (step 2) and a multiple-choice test, comprising an individual word and 4 possible answers (1 correct corresponding pair word and 3 incorrect pair words), approximately 1 hour after waking. This task was digitalized for use in a mobile app, using a chatbot for automatic delivery and data collection. To humanize the chatbot for prototype testing, we named it "Aurora." The Sleep Memories app's development, in-house testing, and evaluation are detailed in Ireland et al [21].

Sleep Memories was designed to collect information about a participant's habitual sleep patterns, through a short survey on opening the app, and deliver the SDM task within 3 hours of the participant's regular sleep time. After the participants respond to the questions about their sleep, the app provides a trial run, enabling participants to become comfortable with its functionalities. In line with Lam et al [6], each word pair is presented to the participants for 10 seconds. For the recall component, the app offers a 10-second window, allowing participants to either type or vocally record their answers, ensuring a user-friendly response period. In the morning, participants are prompted to confirm their bedtime and wake-up

time, which is followed by recall and multiple-choice testing to assess their memory retention.

Qualitative methods were selected for this study, as they are ideally suited to elicit participants' perspectives of Sleep Memories and their experiences with current digital health technologies relative to other methodologies. An interpretative philosophical perspective [22] was selected, as this method allowed us to situate the participants' perspectives within their broader life contexts and allowed for flexibility and adaptability regarding the emergent themes and meanings arising from the data.

Figure 1. Screen captures of the Sleep Memories app. (**A**) Aurora (chatbot) introducing themselves to the participant. (**B**) Aurora asking demographic and habitual sleep pattern questions. (**C**) Aurora giving instructions for the word-pairs task. (**D**) Learning component of the word-pairs task being delivered by the app (**E**) Recall component of the word-pairs task being delivered by the app. (**F**) Multiple-choice testing component of the word-pairs task being delivered by the app.



(A) (B) (C) (D) (E) (F)

Participant Recruitment

Participants were recruited via convenience sampling from the Healthy Brain Ageing (HBA) memory clinic. As described previously [23], the HBA clinic is a specialized memory clinic that accepts referrals from general practitioners and medical specialists and provides comprehensive neuropsychological, medical, and mood assessments for older adults aged >50 years with recent subjective cognitive or mood decline. Exclusion criteria are a Mini-Mental State Examination score<20 [24], intellectual disability, insufficient English proficiency for standardized neuropsychological assessment, history of nonaffective psychiatric disorder (eg, schizophrenia), history of stroke, history of head injury (loss of consciousness>30 minutes) or other neurological disorder (eg, epilepsy), or current substance dependence or abuse. Basic demographic (eg, age, sex, and years of education) and clinical data (eg, Mini-Mental State Examination score and MCI status) collected at the HBA clinic are reported here for descriptive purposes. MCI status was categorized into 3 distinct groups: subjective cognitive impairment (absence of objective cognitive impairment), nonamnestic MCI (objective cognitive impairment in nonmemory domains), and amnestic MCI (defined by objective cognitive impairment in memory domain). Participants were invited to participate in the study via email or phone call. Interested participants were provided with a participant information statement identifying the research team, funding, research aims, and risks and explaining confidentiality, consent, and withdrawal. Once informed consent was obtained, participants were invited to attend 1 of 2 digital focus groups via Zoom (Zoom Video Communications).

Data Collection

Before the focus groups, participants were required to complete a digital survey consisting of an Abbreviated version of the previously validated Mobile Device Proficiency Questionnaire (AMDPQ) [25]. In general, the abbreviated version of the questionnaire captured 7 items on basic proficiency in mobile phone use (eg, charging the phone and typing with the keyboard), 8 items on proficiency in communication (eg, sending emails and messaging), 3 items on data and file storage (eg, transferring files from mobile to computer), 8 items on internet use, and 6 items on troubleshooting and software management (eg, updating and deleting apps). Higher scores on each item indicated greater proficiency in each aspect.

A semistructured focus group guide was developed in collaboration with AL (male, program manager), Peta Mills (female, clinical project officer, PhD), SS (female, postdoctoral research associate), and SLN (female, professor and clinical neuropsychologist with extensive experience in focus groups, workgroups, interviews, and working parties for a variety of purposes such as to develop guidelines or models). The focus groups and questions were structured around three core themes (see Multimedia Appendix 1 for the question guidelines used):

1. Current use of mobile phone apps: informed by prior literature [17,18], this theme focuses on participants' current use of mobile phone apps, particularly for health and sleep-related functions. The theme includes use habits, types of health apps used, and experience with speech recognition features on mobile devices. Related questions were designed to assess participants' familiarity with mobile phone technology, health-related apps, as well as their interactions and experiences with speech recognition features.

2. Interest in completing cognitive tasks through a mobile app: informed by previous studies [18,26], this theme was selected to explore participants' willingness and comfort in using a mobile app for memory assessments. This theme includes the potential benefits and challenges of this approach compared to traditional methods. The aim is to understand the acceptability, motivation, practicality, and barriers or facilitators of adopting mobile apps for cognitive assessments.
3. Feedback on the proposed mobile app: this theme seeks to gather participants' perspectives on various aspects of the proposed mobile app, including the clarity of instructions and tasks, visual appeal, potential distractions, and privacy concerns. Participants were also invited to provide suggestions for enhancing the usability and engagement of the mobile app. The questions specifically targeted feedback on the app's features, such as the chatbot and task length or design, and were developed based on internal testing of the mobile app.

A Microsoft PowerPoint presentation (Multimedia Appendix 2) was used to guide participants through the focus group material. The presentation introduced the concept and significance of SDM, highlighting its potential impairment in aging and cognitive decline and its understudied nature due to current limitations in traditional data collection methods (ie, in a sleep laboratory setting). We showed how our mobile app may be a novel solution to these barriers to facilitate home-based assessments. The presentation also illustrated how participants would interact with Aurora, including the delivery of the word-pairs task, in which the process and structure of the SDM assessment involving learning and recall word pairs were explained.

Data were collected via 2 focus groups between October and November 2022. Each focus group was facilitated by separate researchers: SS, a postdoctoral research associate with 10 years of applied research experience and 4 years of qualitative research experience (interviews and focus groups), and Peta Mills, a clinical project officer with 11 years of applied research experience and 18 months of qualitative research experience (focus groups). At the time of the focus groups, both facilitators were not involved in this research project in any capacity and thus provided an impartial perspective. AL and DI (male, software engineer) attended the focus group sessions to answer participants' logistical and technical questions. Focus groups lasted 60 - 90 minutes and were audio-recorded, transcribed, and de-identified.

Ethical Considerations

This study was approved by The University of Sydney Human Ethics Committee (2022/HE000563) and conducted in accordance with the World Medical Association Declaration of Helsinki. Written informed consent was obtained from all participants prior to any study procedures and given the option to opt out of the study at any given time. Participants were given a A$20 (US $12.75) gift voucher as compensation for their time. All data collected were deidentified to ensure participant privacy and confidentiality.

Data Analysis

For clinical and survey data, frequency and descriptive statistics were computed using SPSS Statistics (version 24.0; IBM Corp). Focus group transcripts were analyzed thematically in NVivo (version 13; Lumivero), a qualitative data analysis computer software package, using the techniques described by Braun and Clarke [27]. The thematic analysis approach described by Braun and Clarke [27] is a systematic, iterative, and inductive process comprising 6 phases. Initially, researchers immerse themselves in the data to discern patterns and potential coding schemes. Subsequently, they generate initial codes to categorize salient features systematically. These codes are then organized into potential themes. Themes undergo rigorous validation to ensure coherence and alignment with the dataset. Once themes are finalized, they are defined, and any hierarchical relationships are identified. Finally, thematic insights are woven into a narrative to address relevant research questions.

As we had identified our 3 core themes during the development of the focus group guide, our thematic analysis was focused on a deeper exploration of these themes. Following Braun and Clarke [27] phases, AL and SS became familiar with the data via repeated reading and individually created initial codes for all transcripts by noting evocative phrases, ideas, and perceptions. Codes were then organized into meaningful groups and sorted into subthemes. Where there was disagreement in coding or theme development, the codes or themes were discussed and refined. To ensure the study conformed to best practice guidelines, we used the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist and met 27 of 32 items [28].

Results

Participants

In total, 19 people from the HBA memory clinic were invited to participate in the focus groups, and 11 of these agreed to participate. Their demographics and clinical information are reported in Table 1. Briefly, the sample on average was aged 68.5 (SD 5.1) years and predominately female (n=7, 64%). The sample also comprised 8 subjective cognitive impairment, 2 nonamnestic MCI, and 1 amnestic MCI. Of these participants, 10 completed the AMDPQ, with 1 participant declined, as they did not have sufficient proficiency to access the digital questionnaire. On average, participants demonstrated relatively high mobile proficiency (see Multimedia Appendix 3 for a full breakdown of participant responses to each individual item).

Table . Participant demographic and clinical information.

	Values
Age (years), mean (SD)	68.5 (5.1)
Sex (male), n (%)	4 (36)
Education (years), mean (SD)	15.1 (2.9)
MMSE[a] score (out of 30), mean (SD)	29.2 (0.9)
Subjective cognitive impairment, n (%)	8 (73)
Nonamnestic mild cognitive impairment, n (%)	2 (18)
Amnestic mild cognitive impairment, n (%)	1 (9)
AMDPQ[b] total score (out of 25), mean (SD)	20.5 (4.6)
AMDPQ—mobile device basic subscale (out of 5), mean (SD)	4.4 (0.6)
AMDPQ—communication subscale (out of 5), mean (SD)	4.0 (1.0)
AMDPQ—data and file storage subscale (out of 5), mean (SD)	3.7 (1.3)
AMDPQ—internet subscale (out of 5), mean (SD)	4.2 (1.1)
AMDPQ—troubleshooting and software management subscale (out of 5), mean (SD)	4.1 (1.1)

[a]MMSE: Mini-Mental State Examination.

[b]AMDPQ: Abbreviated Mobile Device Proficiency Questionnaire.

Thematic Analysis

Overview

The thematic analysis revealed two broad themes relevant to the 3 core focus group themes: (1) context of mobile phone use and (2) shifting from traditional laboratory-based assessments to home assessments. The latter theme should be considered within the framework of the former theme, as shown in Textbox 1. That is, theme 2 considers the challenges and facilitators identified by participants when shifting from laboratory settings to home environments within 2 specific contexts: (1a) participants' existing familiarity and proficiency with mobile phones and apps and (1b) the prevalent stereotypes and preconceived notions concerning aging and technology.

Textbox 1. Breakdown of the overarching themes and subthemes identified in our study by 2 primary domains: "context of mobile phone use" and "shifting from traditional laboratory-based assessments to home assessments."

Context of mobile phone use

- Current phone use
- Aging

Shifting from traditional laboratory-based assessments to home assessments

- Accuracy of data
- Human versus robot
- Accessibility
- Compliance

Context of Mobile Phone Use

Current Phone Use

Participants were asked to share their mobile phone use patterns and specify the types of apps they used. Across participants, mobile phone use varied substantially from the use of basic phone functions, such as phone calls, text messaging, and camera functions, to more complex use involving apps, such as emails, social media, search engines, maps, news, weather, and music. Although minimally used, health and sleep apps were accessed for relaxation and to monitor sleep, heart rate, and physical activity. One participant felt highly proficient in phone use and that the phone was integral to the organization of their day-to-day life: "I would be lost without my mobile phone." Speech-to-text functions were used as needed, particularly when driving and composing text messages, using navigation maps, or playing music. However, it is worth noting that not all participants chose to use this feature.

Aging

Individuals who reported limited phone use perceived their age as a barrier to gaining proficiency with new technologies.

> *I don't think I am alone in this, but it is a real problem when you're a bit older trying to pick up the technology that seems to be intuitive to other people.*

Participants expressed a sense of not keeping pace with technology, which stemmed from their preference for relying on what they were familiar with or their self-perception as old-fashioned. Even when participants successfully used a new technology, they felt the need to use it continuously to avoid losing the knowledge of how to operate it effectively.

> *There is so much learning involved and if you don't use these things more than once in 6 months it's gone.*

Irrespective of the reasons for their limited access to, and use of, technology, a prevailing sentiment among these individuals was a sense of missing out.

> *I have a Samsung phone, one of the latest, and I know there are lots of lovely things for me and I'm not using them. I would use them if they were set up.*

Younger adults were viewed as having an innate familiarity with and competence in using digital tools and, thus, were flagged as potential facilitators of accessing and using technology. Subsequently, participants who did not have a younger adult available to facilitate their use of new technologies felt at a disadvantage. These observations reinforced the prevailing belief that older generations faced a variety of disadvantages with technology.

The insights gathered suggest a spectrum of technology assistance needs among participants. Some individuals may require comprehensive setup support, including step-by-step instructions and human interaction throughout their tech journey, while others might only require minimal assistance. This underscores the importance of adopting 1 of 2 design approaches for apps: either a straightforward and user-friendly design that prioritizes ease of use or, in the case of complex apps, the provision of substantial human support.

These findings carry significant implications for the design of apps intended for home use as opposed to controlled laboratory settings. They serve as a foundation for understanding the crucial factors and context that should inform the app design process when moving apps from the laboratory to the home.

Shifting From Traditional Laboratory-Based Assessments to Home Assessments

Overview

There were several considerations for moving the SDM task from laboratory-based assessments to home assessments raised during focus groups. Participants identified facilitators as well as challenges and proposed potential solutions to overcome these obstacles (Table 2).

Table . Moving sleep-dependent memory from laboratory to home: challenges, solutions, and changes.

Challenge	Potential solution	App updates
Unexpected events during task completion (eg, phone ringing and someone knocking on the door).	Function to pause and recommence task.	The task was not implemented with a pause option because it could impact the results. Instead, Aurora advised participants to find a quiet, distraction-free place at the start of the task. If they could not finish because of unforeseen circumstances, they were offered another chance on a different night.
Difficulty with setting up apps and need help with troubleshooting.	Set up the app for participants and having someone to contact.	Through email, participants can reach a contact person who will provide technical support.
Participants did not want to enter identifying information (eg, names or date of birth).	Using a participant ID rather than identifying information.	We have given the participants an option to enter their participant ID. We now only collect participant initials and not name or date of birth.
Participants flagged that sometimes they ignore push notifications on their phones in the evening.	Option for the phone to ring or a text message.	The app now sends additional notifications to remind the participants to complete the task. Furthermore, the phone now sends short audio notification reminders to participants.

Accuracy of Data

Participants generally regarded the home environment as conducive to achieving higher accuracy levels relative to the laboratory environment, which was described as "ghastly" and distressing. Indeed, because of this perception, participants questioned the reliability of sleep laboratories, primarily because of the unconventional and unfamiliar sleep conditions they imposed. These doubts extended to the execution of the SDM task in the laboratory. Participants posited that conducting the SDM task in the laboratory might cause reduced accuracy compared to conducting the same tasks within the comfort of their homes, where they felt more relaxed as they did not have a researcher watching them and where they could expect a more restful night's sleep. Similarly, there was a recognition that the peculiar sleep laboratory environment might introduce sleep disturbance that may disrupt SDM processes. Therefore, the results of the SDM task completed in a laboratory may not truly reflect an individual's performance. To further illustrate these insights, one participant shared their personal experience with a sleep laboratory:

> *I recently, did a sleep study. It's just not what I consider to be an indicator of anything that's going on. I had a two-and-a-half-hour sleep if that through the whole night strapped into all these other things*

> *as well ... So definitely at home I think you'll get more relaxed, and you get a true picture of what's going on.*

Conversely, having a researcher present and demanding attention could lead to better performance.

> *In the clinic, I was being put on the spot ... "we're not just being casual here, we're doing a trial and I'm seeking your recall results, so you need to pay attention." So being in the clinic I was really on purpose for fear of failure ... At home I'd be far more relaxed and maybe that's a good thing too.*

In the laboratory setting, participants interact with the SDM task by saying their responses aloud. When considering the at-home mobile app version of the SDM task, participants were given a choice between typing their responses into an open text field or vocalizing their answers. Preferences varied, as participants felt that while using verbal responses may be more convenient than typing, the accuracy of verbal responses could be reduced. Nonetheless, being able to choose to use a text or a verbal response resonated positively with participants.

Human Versus Robot

During thematic analysis, some initial disagreements arose among the 2 scorers regarding the classification of human versus robot. AL suggested that the theme "human versus robot" may encompass accessibility as well, while SS suggested creating a separate theme for "accessibility." To resolve this, the team engaged in iterative discussions, referring back to the original transcripts to ensure alignment with participants' narratives. Through this process, a consensus was reached to have "human versus robot" and "accessibility" as 2 separate themes.

The facilitators explained to participants that they would primarily interact with Aurora, a chatbot designed to guide them through the SDM task via written text or verbal instructions. Participants overwhelmingly agreed that any kind of computerized intelligence should be as humanlike and natural as possible, suggesting that a computerized voice might be difficult to engage with.

> *I found it very easy to do it in the lab and the interaction with [researcher name] it was simple to do and it was a softer voice and it is relaxed and it is encouraging. All that subliminal processing of getting the task done.*

Relying solely on a computerized and automated system without human support raised several challenges, and the lack of human interaction was primarily viewed as a drawback. Indeed, participants flagged the importance of human support, both during setup and for ongoing assistance. Participants also identified that the app may struggle to accommodate diverse accents or may lack the mechanism to clarify potential response errors, tasks that were easily accomplished by a human. To assist with these challenges, participants recommended that the app should be capable of detecting potential errors and should offer corrections or alternative suggestions when necessary.

Additionally, participants expressed apprehension around the security of data, given personal information was to be entered directly into the app. There was a hesitation to provide personal information, such as names or dates of birth. Instead, participants proposed using participant IDs only for entry into the app and that other information, such as date of birth, is gathered by the researcher in another way and is linked externally to this ID. However, other participants displayed indifference to this issue, expressing a belief that personal information is already widely accessible: "Everyone knows everything about us anyway."

Accessibility

Participants identified 2 primary barriers that could impact the app's accessibility. These were physical and economic barriers. Several participants expressed concerns regarding age-related vision problems, which can hinder their ability to read or view small text. Participants proposed using larger screens, like tablets, and increasing font sizes might address this barrier. Participant responses from the AMDPQ suggested that certain participants exhibited greater proficiency when using personal computers or tablets than smartphones, indicating the potential value of designing apps with flexibility and accessibility across different platforms.

> *I do I get a bit irritable with the smallest of the screen prefer to do things on the laptop.*

Another physical barrier that was discussed was physical dexterity, which refers to fine motor skills and the ability to move their fingers. Age-related factors, such as tremors and arthritis, may be challenging for older adults to respond to Aurora or complete the SDM task.

> *The dexterity on the fingers may be important on how well someone clicks on the screen ... I might have trouble because I haven't done it on a phone and only do it on a PC.*

One participant raised concerns about whether the app required high bandwidth to use. High bandwidth apps can increase difficulty in accessing and using mobile apps as well as increase costs associated with using the app. Therefore, it is important to optimize apps for low bandwidth.

> *Is your application bandwidth heavy that it's just going to slow things down for those of us who actually don't have a proper wireless connection or a slow data connection and it's just going to make it after a while just unusable nothing.*

Participants noted that the initial setup phase of new technologies (eg, installing and setting up of apps) was either too challenging or time-consuming. Participants recognized that they required assistance or guidance to set up, navigate, and use technology effectively. This support could encompass various aspects, such as learning to use new devices, understanding complex software apps, or troubleshooting technical issues. However, they found that, in general, such support was lacking or insufficient. The consensus among participants strongly supports the idea that offering setup assistance, particularly during the initial stages of technology adoption, would be immensely beneficial. This will help familiarize participants with the mobile app, save time, and alleviate stress for older adults. Any kind of automation was also supported to avoid repeated data entry:

> *I don't need to put that same information in every time that I'm doing this test because this would be 5 - 10 minutes every day just putting in the same information ... when you are actually setting up the app on your phone is when you actually key that data in and then you don't have to do it again*

Compliance

The facilitators explained to the participants that the home-based completion of the chatbot would require dedicating approximately 40 minutes to the learning component, followed by a subsequent delay of 20 - 30 minutes and another 10 minutes to complete the task. While participants expressed comfort with the overall duration of the task, the true challenge was quarantining the time to do this task.

> *I think anything more than 40 minutes it may be burdensome. I do think 40 minutes is the top. The challenge is disruptions and to really quarantine the time and be strong about that.*

Participants proposed a solution, emphasizing the necessity of proactively earmarking dedicated time in advance to ensure the completion of the task. Having a large time window during which participants could complete the task would provide the flexibility needed for accommodating various schedules and preferences of participants.

> *I would quarantine in the time then for the next day and prepare myself because I don't have any other commitments or pressures on my time.*

Alternatively, participants suggested introducing a pause feature to mitigate any potential disruptions caused by unforeseen events, such as incoming phone calls or unexpected visitors. However, it is crucial to acknowledge that the introduction of a pause feature could potentially influence the integrity of the test. While these features may be convenient for the participants, this will introduce confounds to the results. Subsequently, a pause button was not implemented in Aurora.

> *I always find a pause button is important. For example: if I am starting the memory test or memory exercise right now and someone knock the door ... I have the opportunity that I can come back to that when I need to do something.*

Another challenge identified was participants occasionally forgetting to complete the task. The facilitators explained to the participants that there are existing notifications within the app to remind the participants to do the task in the evening and in the morning. It was suggested that increasing the frequency of notifications could be a viable solution. Moreover, participants highlighted a tendency to ignore push notifications, particularly in the evening. To address this, participants proposed implementing alternative reminder methods, such as phone calls or text messages. Participants identified that their interest and motivation to complete the memory task were closely tied to the app's clarity. The consensus was that instructions and information presented in an easily digestible manner would keep participants' engagement. It was evident that it is important to prioritize clear and straightforward instructions that would lead to sustained participant engagement, especially when participants are asked to complete the task on multiple occasions.

> *Okay, if you wanted me to use it repeatedly then it's got to be clear and simple and easy to access if you're talking about once every three months and that's ok I'll come back anytime.*

Participants nearly unanimously highlighted the critical role of feedback on their task performance. They underscored the importance of receiving real-time feedback to gauge their progress and enable a sense of accomplishment and motivation to improve and continue to complete the task for multiple assessments.

> *I don't want money but I do want feedback. I am involved in a lot of medical research I am surprised I don't find out about the results.*

Even participants who did not think that feedback was necessary thought it would be a welcomed supplementary feature, as it would motivate them to continue using the app. Subsequently, Sleep Memories integrated informative feedback to enhance user engagement and promote continued motivation. For others, the motivation was intrinsic (ie, feedback not needed), and people would remain motivated to contribute to science. Participants trusted that scientists knew what they were doing and would follow their guidance because they are experts in the field.

> *... The work that this unit does is so important. Just thanking them for their time and delivering, any material or rewards I find it degrading. I just want to donate my data to be used for the greater good. There is nothing you can give me to make it better.*

Discussion

Principal Findings

This study aimed to investigate the perspectives of older adults both with subjective or objective cognitive impairment regarding barriers and facilitators to using an at-home assessment of SDM through a novel app named Sleep Memories. Through co-design methodologies, we assessed the initial prototype of the mobile app to incorporate user feedback into a revised version suitable for clinical testing.

We found that digitalization and technological advancements can both facilitate and hinder everyday life for older adults. This phenomenon is reflected in the term Janus-faced technology [29]. According to this concept, the successful integration of technology can substantially improve daily activities for older adults. However, barriers to use can evoke feelings of alienation and disconnection from the digital world. Understanding these barriers is essential to minimize these negative feelings and to, instead, develop accessible and user-friendly technology solutions for older adults, particularly those with cognitive impairments.

One of the key findings of our study was the importance of app flexibility in accommodating users' diverse needs and preferences and in challenging barriers. While some individuals reported requiring high-level assistance, others indicated that

they could navigate the app independently with minimal support. A recent systematic review highlighted the critical importance of tailoring technology to meet the specific needs and preferences of older adults with MCI and dementia [30]. Below we examine each of the facilitators and barriers we identified and explore corresponding, flexible solutions in light of existing research.

Comparison to Prior Work

The "first night effect" is a confounding variable in sleep research where participants' sleep is disrupted due to being in an unfamiliar environment. This effect has been studied in prior studies. For instance, one study [31] demonstrated its potential to impact sleep disorder diagnoses. The benefit of home-based testing is that participants can sleep in their usual environment, removing the need for multiple night sleep laboratory testing. Furthermore, the SDM data collected are not confounded by the first night effect and more accurately represent a participant's true abilities than traditional SDM testing.

Within the discussion of voice responses, our findings align with prior work, with concerns from older adults regarding the accuracy of audio (eg, the chatbot's ability to accurately capture verbal responses) and text (eg, spelling mistakes) inputs [32]. In response to these concerns, Sleep Memories has implemented a dual-layered validation process. First, an autocorrect function is used to rectify minor spelling inaccuracies automatically. Second, when the system is uncertain of the response, a request for confirmation by a researcher is made. The implementation and availability of a voice response option increase the accessibility of the mobile app for those who may have age-related declines or disabilities [32].

To increase the accessibility of mobile apps for older adults, it is important to navigate both technological and socioeconomic challenges. Our focus groups highlighted concerns like high bandwidth and data consumption, which could impede the functionality and availability of health apps such as "Sleep Memories" for older adults. Furthermore, older adults may have age-related physical limitations, such as arthritis or hand pain, which may impact their interaction with mobile devices. To address the high bandwidth and data consumption, we allow the app to be used without an internet connection, and data are uploaded when connectivity is available. As previously mentioned, the implementation of voice responses can be a solution to enable those with physical limitations to use the app.

Participants raised concerns about forgetting to complete the task, particularly for those who might overlook setting personal reminders. This aligns with previous research in the context of medication compliance, where it was observed that a single reminder was often insufficient to prompt action, suggesting that a backup notification could serve as a viable solution [33]. A solution in our context that was proposed was to increase the frequency and elevate the distinctiveness of reminder notifications (eg, for Sleep Memories, we implemented a unique auditory notification, specifically a vocal prompt stating "it is now time to complete your memory task").

Similar to a previous study [34], we found that incorporating feedback mechanisms, like scores, within eHealth tools is imperative to foster user engagement and empowerment. These features motivate users by providing immediate, tangible feedback on their performance. However, it is essential for app developers and researchers to communicate the context and limitations of their scores, particularly when there is a lack of established normative data. Without proper context, participants may misinterpret their scores, leading to unnecessary anxiety or concerns.

Our study revealed that conducting tasks within a home setting may introduce various distractions. This was aligned with prior research on the environmental interference of unsupervised home cognitive assessments [26]. To mitigate this challenge, 2 solutions emerged during our focus group discussions. The first involves incentivizing participants' engagement through targeted motivation strategies, such as feedback, to maintain focus despite potential distractions. The second suggests the need for flexibility in task design, allowing participants to pause and resume activities, accommodating the unpredictability of the home environment. However, this approach must be balanced against the risk of compromising data integrity. For instance, in "Sleep Memories," participants can pause and take short breaks between trials without permitting interruptions during the trials themselves to safeguard data continuity and accuracy.

Limitations

Our study comprised older adults who had relatively high mobile proficiency at least in the subdomains that we examined, so our findings may not be generalizable to older adults who have a low mobile proficiency. Another limitation is that our sample had less representation from individuals with MCI (n=3), which limited our ability to detect or uncover differences in responses related to cognitive impairment. Future research is necessary to explore this important area by recruiting a larger cohort of individuals with MCI. Similarly, our participants generally scored highly on the AMDPQ, meaning that they showed decent proficiency in mobile phone use. Whether these findings apply to older adults with low proficiency in mobile phones needs further investigation. Nonetheless, our study is one of the first to examine the potential barriers and facilitators to the use of mobile device apps, in particular to complete memory tasks, in those with cognitive impairment.

Future Directions

In addition to the app modifications mentioned earlier, our findings provide valuable insights to meet the needs and preferences of our older adult users. When deploying Sleep Memories, we will first conduct a needs assessment to assess the ability to use existing and new technology as well as the user's circumstances (eg, Do they have time to complete the task in the app without distraction?) and environment (eg, Do they have access to a technology-savvy person in their network?). Based on this assessment, we will offer users various levels of support such as the ability for a phone call by a researcher or various notification options.

Conclusions

Our study demonstrates that while there are barriers related to accessibility, usability, and data integrity to using health-related phone apps in this population, there are important facilitators

that can be implemented in phone apps to create a flexible and inclusive digital health tool. Furthermore, we have shown evidence on the potential feasibility of the "Sleep Memories" app to enable the collection of SDM data within a user's natural environment.

Acknowledgments

The authors would like to extend their deepest gratitude to all the participants who dedicated their time and shared their experiences with them. The authors also wish to acknowledge Peta Mills for her assistance with their focus groups. Finally, the authors thank their clinical and research team at the Healthy Brain Ageing clinic. This study was supported by the National Health and Medical Research Council (NHMRC) Centre of Research Excellence to Optimise Sleep in Brain Ageing and Neurodegeneration (CogSleep), the Commonwealth Scientific and Industrial Research Organisation, and an NHMRC Fellowship awarded to SLN (GNT1135639) and AD (GNT2008001).

Data Availability

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

Conflicts of Interest

SLN has received consulting fees from Eisai and Roche pharmaceuticals. All other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Discussion guide for focus groups.
[DOCX File, 19 KB - aging_v8i1e68147_app1.docx]

Multimedia Appendix 2
Chatbot for word pairs.
[DOCX File, 4460 KB - aging_v8i1e68147_app2.docx]

Multimedia Appendix 3
Abbreviated Mobile Device Proficiency Questionnaire.
[DOCX File, 26 KB - aging_v8i1e68147_app3.docx]

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2023-11-01]
2. Förstl H, Kurz A. Clinical features of Alzheimer's disease. Eur Arch Psychiatry Clin Neurosci 1999;249(6):288-290. [doi: 10.1007/s004060050101] [Medline: 10653284]
3. Zuniga KE, Mackenzie MJ, Kramer A, McAuley E. Subjective memory impairment and well-being in community-dwelling older adults. Psychogeriatrics 2016 Jan;16(1):20-26. [doi: 10.1111/psyg.12112] [Medline: 25737426]
4. Stickgold R, Walker MP. Sleep-dependent memory triage: evolving generalization through selective processing. Nat Neurosci 2013 Feb;16(2):139-145. [doi: 10.1038/nn.3303] [Medline: 23354387]
5. Feld GB, Weis PP, Born J. The limited capacity of sleep-dependent memory consolidation. Front Psychol 2016;7:1368. [doi: 10.3389/fpsyg.2016.01368] [Medline: 27679589]
6. Lam A, Haroutonian C, Grummitt L, et al. Sleep-dependent memory in older people with and without MCI: the relevance of sleep microarchitecture, OSA, hippocampal subfields, and episodic memory. Cereb Cortex 2021 May 10;31(6):2993-3005. [doi: 10.1093/cercor/bhaa406] [Medline: 33565576]
7. Lam A, Kong S, Naismith SL. Recent advances in understanding of sleep disorders and disturbances for dementia risk and prevention. Curr Opin Psychiatry 2024 Mar 1;37(2):94-100. [doi: 10.1097/YCO.0000000000000921] [Medline: 38226546]
8. Orlando IF, O'Callaghan C, Lam A, et al. Sleep spindle architecture associated with distinct clinical phenotypes in older adults at risk for dementia. Mol Psychiatry 2024 Feb;29(2):402-411. [doi: 10.1038/s41380-023-02335-1] [Medline: 38052981]
9. Lam AKF, Carrick J, Kao CH, et al. Electroencephalographic slowing during REM sleep in older adults with subjective cognitive impairment and mild cognitive impairment. Sleep 2024 Jun 13;47(6):zsae051. [doi: 10.1093/sleep/zsae051] [Medline: 38394454]
10. Djonlagic IE, Guo M, Igue M, Kishore D, Stickgold R, Malhotra A. Continuous positive airway pressure restores declarative memory deficit in obstructive sleep apnea. Am J Respir Crit Care Med 2021 May 1;203(9):1188-1190. [doi: 10.1164/rccm.202011-4253LE] [Medline: 33347378]

11. Ketz N, Jones AP, Bryant NB, Clark VP, Pilly PK. Closed-loop slow-wave tACS improves sleep-dependent long-term memory generalization by modulating endogenous oscillations. J Neurosci 2018 Aug 15;38(33):7314-7326. [doi: 10.1523/JNEUROSCI.0273-18.2018] [Medline: 30037830]
12. Papalambros NA, Santostasi G, Malkani RG, et al. Acoustic enhancement of sleep slow oscillations and concomitant memory improvement in older adults. Front Hum Neurosci 2017;11:109. [doi: 10.3389/fnhum.2017.00109] [Medline: 28337134]
13. Gui WJ, Li HJ, Guo YH, Peng P, Lei X, Yu J. Age-related differences in sleep-based memory consolidation: a meta-analysis. Neuropsychologia 2017 Mar;97:46-55. [doi: 10.1016/j.neuropsychologia.2017.02.001] [Medline: 28161367]
14. Harrison EI, Roth RH, Lobo JM, et al. Sleep time and efficiency in patients undergoing laboratory-based polysomnography. J Clin Sleep Med 2021 Aug 1;17(8):1591-1598. [doi: 10.5664/jcsm.9252] [Medline: 33739259]
15. Mace RA, Mattos MK, Vranceanu AM. Older adults can use technology: why healthcare professionals must overcome ageism in digital health. Transl Behav Med 2022 Dec 30;12(12):1102-1105. [doi: 10.1093/tbm/ibac070] [Medline: 36073770]
16. LaMonica HM, Milton A, Braunstein K, et al. Technology-enabled solutions for Australian mental health services reform: impact evaluation. JMIR Form Res 2020 Nov 19;4(11):e18759. [doi: 10.2196/18759] [Medline: 33211024]
17. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
18. Wilson SA, Byrne P, Rodgers SE, Maden M. A systematic review of smartphone and tablet use by older adults with and without cognitive impairment. Innov Aging 2022;6(2):igac002. [doi: 10.1093/geroni/igac002] [Medline: 35243008]
19. Andrews NE, Ireland D, Vijayakumar P, et al. Acceptability of a pain history assessment and education chatbot (Dolores) across age groups in populations with chronic pain: development and pilot testing. JMIR Form Res 2023 Oct 6;7:e47267. [doi: 10.2196/47267] [Medline: 37801342]
20. Piau A, Crissey R, Brechemier D, Balardy L, Nourhashemi F. A smartphone chatbot application to optimize monitoring of older patients with cancer. Int J Med Inform 2019 Aug;128:18-23. [doi: 10.1016/j.ijmedinf.2019.05.013] [Medline: 31160007]
21. Ireland D, Lam A, Bradford D, Naismith S. A new role for chatbots: automation of a sleep-dependent memory task. Stud Health Technol Inform 2024 Jan 25;310:519-523. [doi: 10.3233/SHTI231019] [Medline: 38269863]
22. Spencer R, Pryce JM, Walsh J. Philosophical approaches to qualitative research. In: The Oxford Handbook of Qualitative Research: Oxford University Press; 2014:81-98.
23. Duffy SL, Lagopoulos J, Hickie IB, et al. Glutathione relates to neuropsychological functioning in mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):67-75. [doi: 10.1016/j.jalz.2013.01.005] [Medline: 23688577]
24. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
25. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the Mobile Device Proficiency Questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
26. Madero EN, Anderson J, Bott NT, et al. Environmental distractions during unsupervised remote digital cognitive assessment. J Prev Alzheimers Dis 2021;8(3):263-266. [doi: 10.14283/jpad.2021.9] [Medline: 34101782]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Tong A, Sainsbury P, Craig J. COnsolidated criteria for REporting Qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042]
29. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. "These devices have not been made for older people's needs"—older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
30. Holthe T, Halvorsrud L, Karterud D, Hoel KA, Lund A. Usability and acceptability of technology for community-dwelling older adults with mild cognitive impairment and dementia: a systematic literature review. Clin Interv Aging 2018;13:863-886. [doi: 10.2147/CIA.S154717] [Medline: 29765211]
31. Newell J, Mairesse O, Verbanck P, Neu D. Is a one-night stay in the lab really enough to conclude? First-night effect and night-to-night variability in polysomnographic recordings among different clinical population samples. Psychiatry Res 2012 Dec 30;200(2-3):795-801. [doi: 10.1016/j.psychres.2012.07.045] [Medline: 22901399]
32. Werner L, Huang G, Pitts BJ. Smart speech systems: a focus group study on older adult user and non-user perceptions of speech interfaces. Int J Hum-Comput Interact 2023 Mar 16;39(5):1149-1161. [doi: 10.1080/10447318.2022.2050541]
33. Stawarz K, Cox A, Blandford A. Don't forget your pill!: designing effective medication reminder apps that support users' daily routines. Presented at: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 26 to May 1, 2014; Toronto, Ontario, Canada. [doi: 10.1145/2556288.2557079]
34. Thompson HJ, Demiris G, Rue T, et al. A holistic approach to assess older adults' wellness using e-health technologies. Telemed J E Health 2011 Dec;17(10):794-800. [doi: 10.1089/tmj.2011.0059] [Medline: 22011052]

Abbreviations

AMDPQ: Abbreviated Mobile Device Proficiency Questionnaire
COREQ: Consolidated Criteria for Reporting Qualitative Research
HBA: Healthy Brain Ageing
MCI: mild cognitive impairment
SDM : sleep-dependent memory

Edited by M Mardini; submitted 29.10.24; peer-reviewed by E Okoyeocha, M Velazquez, RS Gomaa Mahmoud; revised version received 12.02.25; accepted 18.03.25; published 28.04.25.

Please cite as:
Lam A, Simonetti S, D'Rozario A, Ireland D, Bradford D, Fripp J, Naismith SL
Perceptions of the Use of Mobile Apps to Assess Sleep-Dependent Memory in Older Adults With Subjective and Objective Cognitive Impairment: Focus Group Approach
JMIR Aging 2025;8:e68147
URL: https://aging.jmir.org/2025/1/e68147
doi:10.2196/68147

Relationship Between Within-Session Digital Motor Skill Acquisition and Alzheimer Disease Risk Factors Among the MindCrowd Cohort: Cross-Sectional Descriptive Study

Andrew Hooyman[1], PhD; Matt J Huentelman[2], PhD; Matt De Both[2], BS; Lee Ryan[3], PhD; Kevin Duff[4], PhD; Sydney Y Schaefer[5], PhD

[1]Department of Physical Therapy, Chapman University, 9401 Jeronimo Rd, Irvine, CA, United States

[2]Division of Early Detection and Prevention, The Translational Genomics Research Institute, Phoenix, AZ, United States

[3]Department of Psychology, The University of Arizona, Tucson, AZ, United States

[4]Department of Neurology, Oregon Health and Science University, Portland, OR, United States

[5]School of Biological Health Systems Engineering, Arizona State University, Tempe, AZ, United States

Corresponding Author:
Andrew Hooyman, PhD
Department of Physical Therapy, Chapman University, 9401 Jeronimo Rd, Irvine, CA, United States

Abstract

Background: Previous research has shown that in-lab motor skill acquisition (supervised by an experimenter) is sensitive to biomarkers of Alzheimer disease (AD). However, remote unsupervised screening of AD risk through a skill-based task via the web has the potential to sample a wider and more diverse pool of individuals at scale.

Objective: The purpose of this study was to examine a web-based motor skill game ("Super G") and its sensitivity to risk factors of AD (eg, age, sex, *APOE* ε4 carrier status, and verbal learning deficits).

Methods: Emails were sent to 662 previous MindCrowd participants who had agreed to be contacted for future research and have their *APOE* ε4 carrier status recorded and those who were at least 45 years of age or older. Participants who chose to participate were redirected to the Super G site where they completed the Super G task using their personal computer remotely and unsupervised. Once completed, different Super G variables were derived. Linear and logistic multivariable regression was used to examine the relationship between available AD risk factors (age, sex, *APOE* ε4 carrier status, and verbal learning) and distinct Super G performance metrics.

Results: Fifty-four participants (~8% response rate) from the MindCrowd web-based cohort (mean age of 62.39 years; 39 females; and 23 *APOE* ε4 carriers) completed 75 trials of Super G. Results show that Super G performance was significantly associated with each of the targeted risk factors. Specifically, slower Super G response time was associated with being an *APOE* ε4 carrier (odds ratio 0.12, 95% CI 0.02-0.44; *P*=.006), greater Super G time in target (TinT) was associated with being male (odds ratio 32.03, 95% CI 3.74-1192,61; *P*=.01), and lower Super G TinT was associated with greater age (β −3.97, 95% CI −6.64 to −1.30; *P*=.005). Furthermore, a sex-by-TinT interaction demonstrated a differential relationship between Super G TinT and verbal learning depending on sex ($\beta_{male:TinT}$ 6.77, 95% CI 0.34-13.19; *P*=.04).

Conclusions: This experiment demonstrated that this web-based game, Super G, has the potential to be a skill-based digital biomarker for screening of AD risk on a large scale with relatively limited resources.

(JMIR Aging 2025;8:e67298) doi:10.2196/67298

KEYWORDS

digital health technology; web-based assessment; aging; APOE; motor skills; sensitivity; risk factors; adults; older adults

Introduction

Since the number of cases of Alzheimer disease (AD) is expected to double in the next 2 decades [1], there is an urgent need for widespread screening of older adults for their individual AD risk profile, which has implications for clinical care and research. Current options, such as positron emission tomography for measuring tau and beta amyloid pathology, tend to be expensive and invasive and require advanced imaging facilities [2]. Blood-based biomarkers may make screening for AD more affordable and less invasive, but their validity and standardization are still being established at this time [3,4]. Cognitive screening tests can indicate deficits, but these largely rely upon in-person administration from a trained clinician and may have reduced sensitivity in identifying individuals at high risk of AD in the earliest stages [5]. Thus, it is important to

identify measures that are sensitive to AD risk and yet can also be delivered accurately, easily, and directly to patients and prospective participants in clinical trials and research studies.

Digital biomarkers have great potential to meet the need for accessible and remote testing of AD risk. Broadly speaking, digital biomarkers are measurable indicators of health or disease collected from a digital device or through digital means. Some digital biomarkers of AD include finger tapping [5,6], repeated cognitive assessment on a Wi-Fi–enabled device [7], recorded speech [8], and digital clock drawing [9], which have all been shown to be sensitive to cognitive impairment. These measures have also been associated with disease status, differentiating between cognitively intact versus mild cognitive impairment (MCI) and MCI versus dementia to some degree. However, these examples have overall low sensitivity to risk factors of AD, such as age [10], sex [11], apolipoprotein-E (*APOE*) ε4 carrier status [12,13], and brain amyloid [9]. Considering this, another potential digital biomarker of AD could be motor skill. Motor skill acquisition is the within-session improvement in a motor skill as a function of practice [14]. Previous in-lab studies have associated motor skill deficits with *APOE* carrier status, hippocampal atrophy, functional decline, and amyloid deposition among people diagnosed with amnestic MCI [15-18]. This task can also be collected remotely [19,20], which would allow for a wider and more diverse sample of individuals at potential risk for AD and can more easily facilitate longitudinal testing as desired.

This study developed a web-based tool for assessing motor skill performance called Super G [21], which can be reliably played unsupervised on the web [20] regardless of device type and without downloading any app. Specifically, the objective of this study was to examine whether Super G performance was individually related to known risk factors of AD (eg, age, sex, *APOE* ε4 carrier status, and verbal learning). Based on prior motor skill studies in AD, it was hypothesized that Super G performance would be negatively associated with each risk factor.

Methods

Study Design

This was a cross-sectional descriptive study that examined within-session performance characteristics from a web-based motor skill task (Super G) and their association with AD risk factors among adults recruited from the MindCrowd study.

Participants

Participants were recruited in May of 2021 through MindCrowd, a web-based research study launched in 2013 to crowdsource demographic, medical history, lifestyle, and cognitive data to identify risk factors of AD [22,23]. Emails were sent through MindCrowd to a subset of 662 individuals who met our inclusion or exclusion criteria and were older than 45 years, who had previously provided a dried blood spot or saliva sample for *APOE* genotyping (see section “AD Risk Factors and Other Participant Characteristics From MindCrowd” for details) and had provided consent to be contacted for future studies. A hyperlink was included in the email that directed individuals to the Super G game website, on which participants digitally provided consent (approved by the Arizona State University institutional review board study no. 13081). Of the 662 individuals emailed, 54 participants (age: mean 62.39, SD 7.4 years; female: n=39) appropriately registered and completed all 75 trials of the game, equating to an 8.1% response rate. The mean (SD) time between MindCrowd data collection (specifically verbal learning, see section “AD Risk Factors and Other Participant Characteristics From MindCrowd” for details) and Super G data collection was 5.9 (1.4) years.

Super G

The Super G game was developed in Unity 5.3.1 and is hosted on Hostinger. Thus, participants were not required to download an app or program to their device. Super G was developed as a gamified version of a seminal motor skill paradigm [24] and has been validated against the original version [21]. The goal of Super G is to help an astronaut explore as many planets as possible within the game’s solar system. There are 16 planets to visit but only 75 attempts to reach them all. Participants use the left and right arrow keys on their keyboard to move the astronaut onto a planet. However, the game uses a rate control mechanism that may not be immediately apparent to participants [21]. Specifically, pressing the right arrow key applies a constant positive force to move the astronaut toward the planet, while pressing the left arrow key applies an equal negative force away from the planet. Since the virtual environment lacks gravity or drag, any force applied will result in a constant velocity until a negative force is applied to slow it down. Thus, participants must learn to apply negative force at the right time and for the right duration to land the astronaut on each planet.

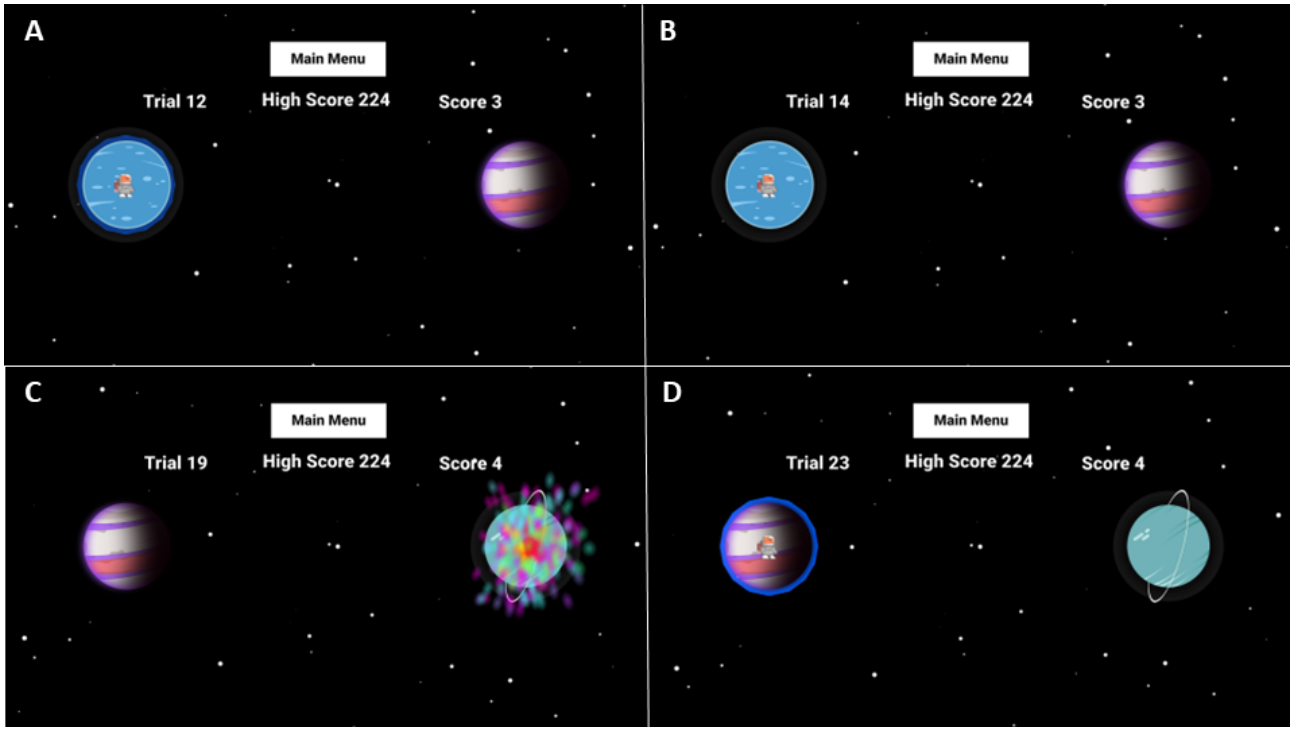
Each trial in Super G begins with the astronaut positioned on the left side of the screen on the initial start planet. The target planet is located on the opposite side of the screen to the right (Figure 1A). The trial lasts 4.5 seconds, but the astronaut cannot leave the start planet until 1.5 seconds have elapsed, as indicated by the disappearance of the blue atmosphere around the start planet (Figure 1B). If the astronaut leaves too early, the trial resets and the astronaut is returned to the initial position on the start planet. Once the blue atmosphere disappears, the participant has 3 seconds to land the astronaut on the target planet. To achieve a successful landing, the astronaut must stay within the boundary of the target planet for 1 continuous second. If successful, a reward tone plays, and fireworks erupt from the target planet (Figure 1C). After a successful landing, Super G repositions at the start position, and the previously landed planet becomes the start planet, with a new target planet appearing in its place (Figure 1D). In the event of a failed landing, the planets for the next trial remain the same, and astronaut reappears at the start position. All 75 trials are completed within a single session for a total session length of approximately 6 minutes.

Cursor position and key press data from each trial of Super G were collected at 100 Hz, along with high scores, time and date of each trial. From these data, four performance variables were calculated. First, scaling ratio (SR) represents the ratio of negative force (duration of left button press) applied to Super G over positive force (duration of right button press), where values equal to 1 indicate equal scaling of forces. Second, time

of reversal (TR) is a measure of how well participants execute the timing of their movements by identifying when the left arrow key press occurs during the trial, where higher values are reversals that occur later in the trial. Successful trials require equal scaling of forces (ie, SR close to or equal to 1) and a reversal timed at the midway point of the trial, although we note that there can be trade-offs between the SR and TR to still allow for some degree of success in the task. Third, total time in the target (TinT) planet is the time that the cursor stayed within the target planet, with higher values representing better performance. Fourth, response time (RT) is a measure of how soon the cursor left the home planet after the disappearance of the start planet's atmosphere, and lower values equate to a faster RT.

Although these performance metrics do strongly correlate with one another, each represents distinct phases of individual motor skill acquisition, as it is possible to have a fast RT with a low TinT if either timing of cursor reversal or scaling of forces is not well executed. Overall, average TinT across all 75 trials is the primary measure of performance since it directly represents the task goal and performance of the task as it relates to execution of SR and TR. Average RT across all 75 trials was considered as a secondary measure of performance because it represents the earlier stage of skill acquisition as participants first need to anticipate when to exit the start planet prior to execution of their acquired movement strategy. Average TinT and RT across the 75 trials are used as the Super G metrics of individual performance due to our prior work [21], which demonstrated that average TinT and RT well describe individual within-session change in performance and delayed retention.

Figure 1. Each panel represents the different phases a participant may experience during Super G play. (A) The astronaut, Super G, spawns on the start planet at the beginning of each trial. There is a blue atmosphere around the planet that signals to the participant that Super G cannot leave yet. (B) Once the blue atmosphere disappears, then Super G can leave the start planet and attempt to land on the target planet. (C) If Super G successfully stays within the target planet boundary for 1 continuous second, then a reward tone plays and fireworks erupt from the planet. (D) Subsequent trials then render the previous target planet as the new start planet and a new target planet is put in place.



AD Risk Factors and Other Participant Characteristics From MindCrowd

Super G data were harmonized with MindCrowd through merging of hashed email addresses linked to MindCrowd data and used as the login for the Super G game. Super G variables were then merged with participant age, sex, level of education, verbal learning (measured via paired associates learning [PAL]) score [25], simple visual reaction time (svRT), and *APOE* genotype data, which were available in MindCrowd. The PAL is a verbal learning task that measures the ability to remember the associations between different word pairs. Specifically, participants are visually presented with 12 word pairs, with each word pair presented separately and at 2-second intervals. Participants were then presented with the first word of each pair and then used their keyboard to type the missing word. This procedure was repeated for 2 additional trials. The maximum score on the PAL is 36 (12 words across 3 trials). The svRT task measured the median reaction time across 6 trials, in which participants had to press any keyboard button as soon as a target stimulus appeared on their screen, with reaction time for each trial recorded in milliseconds. Median svRT was used due to the skewness of the distribution of reaction time in each participant's individual data (skewness svRT median 1.2 vs skewness svRT mean 4.1). The svRT test was used in this study as a control variable only. Participants were classified as *APOE* ε4 carriers or noncarriers. Carriers were defined as individuals who had either 1 or 2 copies of the ε4 allele, and noncarriers were defined as individuals with 0 copies of the ε4 allele. This was based on prior genotyping from biospecimen collection via

self-administered saliva or dried blood spot kits that were mailed to the participants by MindCrowd and then processed by the Translational Genomics Research Institute. Details regarding biospecimen collection and genotyping can be found here [26]. Complete visualization of *APOE* ε4 carrier status with svRT, PAL, and age can be viewed in Figure S1 in Multimedia Appendix 1.

Statistical Analysis

Wilcoxon rank sum tests were used due to a difference in sample sizes between *APOE* ε4 carriers (n=23) and noncarriers (n=31), since uneven sample sizes may lead to inaccurate and disproportionate estimates of variance of each group and violate assumptions of parametric testing. This determined whether there were differences in age, PAL score, svRT, and Super G performance between the 2 groups. Chi-square tests were used to determine whether there was a difference in proportion based on sex, race, education, and ethnicity between carriers and noncarriers. To provide broader context of the verbal learning (PAL) and reaction time (svRT) of this sample, percentile scores were calculated for each participant adjusted for their age, sex, and level of education. This allowed for better contextualization of the relative performance on each measure between carriers and noncarriers relative to the entire MindCrowd cohort (ie, if participants in this study are under- or overperformers compared with what would be expected to a random sample across the entire cohort).

Separate linear and logistic multivariable regression analyses, depending on outcome variable type, were used to test the relationship between Super G performance and the dependent variables, that is, AD risk factors. Specifically, different models were constructed to identify which Super performance metrics are related to specific AD risk factors, while controlling for possible confounding effects of the other AD risk factors and participant characteristics. Multivariable logistic regression was used for the dependent variable of *APOE* ε4 carrier status (where carriers were coded as “true” and noncarriers as “false”), along with the primary and secondary measures of Super G performance (TinT and RT), and control variables of age, sex, PAL score, hour of day Super G played, level of education, and svRT. This approach controlled for the other factors and was repeated with sex as the dependent variable (whereby male was coded as “true” and female as “false”) while switching *APOE* ε4 carrier status to a control variable. Multivariable regression was used when the dependent variable was PAL score, with a primary independent variable of mean Super G performance (TinT, RT, SR, or TR), and control variables of age, sex, *APOE* ε4 carrier status, hour of day Super G played, level of education, and svRT. The same approach was repeated with age as the dependent variable while switching PAL score to a control variable. In addition, to control for the potential delay between initial PAL scores and Super G measurement, PAL-adjusted scores were also generated and analyzed. Individual PAL scores were adjusted based on the amount of time between their PAL and Super G measurement, given previously reported expected decline in PAL based on age reported by Talboom and colleagues [26]. This resulted in an expected 0.2-point decline in PAL for every year between the initial PAL and current Super G measurement. The formulation of each model can be viewed in the following equations:

1. *APOE* ε4 carrier status ~ Super G Performance (TinT or RT) + age +sex + hour played + education +svRT + PAL (Logistic regression)
2. Sex ~ Super G Performance (TinT or RT) + age + hour played + education + svRT + PAL + *APOE* ε4 carrier status (Logistic regression)
3. Age ~ Super G Performance (TinT or RT) + sex + hour played + education + svRT + PAL + *APOE* ε4 carrier status (Linear regression)
4. PAL ~ Super G Performance (TinT or RT) + age + sex + hour played + education + svRT + *APOE* ε4 carrier status (Linear regression)
5. PAL adjusted ~ Super G Performance (TinT or RT) + age + sex + hour played + education + svRT + *APOE* ε4 carrier status (Linear regression)

Participants could play Super G at any time throughout the day, the variable of hour played (measured with a 24-hour clock rather than a 12-hour clock) was transformed using a cosine function to ensure that adjacent hours 0 and 23 were close together. This is visualized in Figure S2 in Multimedia Appendix 1. All numeric variables were standardized (age, svRT, PAL, hour played, and Super G Performance) to be centered at their respective mean and divided by their respective SD. Thus, a 1-unit change in the results of these variables with respect to all reported odds ratios (ORs) and beta coefficients represents a 1 SD change in either the outcome or the performance metric. This allowed for better relative comparison between all independent variables within each model. To detect multicollinearity, the variance inflation factor was calculated, and any variable with a variance inflation factor of >5 was removed. Outliers were identified using Cook’s distance and removed if their distance was >1.

Ethical Considerations

The study protocol was reviewed and approved by the office of Research Integrity and Assurance at Arizona State University (approval number STUDY00013081). To participate, participants needed to provide digital informed consent. All data collected from this experiment were deidentified for privacy and confidentiality. Participants were not compensated for their time in this study.

Results

Participant Characteristics

Overall, there were 23 *APOE* ε4 carriers (20 heterozygotes and 3 homozygotes) and 31 noncarriers. Between-group comparisons (carriers vs noncarriers) using the Wilcoxon rank sum test demonstrated that groups did differ by age (W=471; *P*=.04), with noncarriers being an average of 4.3 years older than carriers (carriers=59.9 years, noncarriers=64.2 years). For all other control variables, there were no observed group differences (Table 1).

Table . Genetic, demographic, cognitive, motor, and Super G performance data between *APOE* ε4 carries and noncarriers.

	APOE ε4 carriers	*APOE* ε4 noncarriers	Test statistic (*df*)	*P* value
Participants, n	23	31	—[a]	—
APOE alleles (X/X), n			—	—
2/2	—	0		
2/3	—	2		
3/3	—	29		
2/4	4	—		
3/4	16	—		
4/4	3	—		
Age (years), mean (SD)	59.9 (7.4)	64.2 (6.9)	*W*=471 (—)	.045[b]
Sex (male/female)	6/17	9/22	χ^2=0 (1)	>.99
Race			χ^2=0 (1)	>.99
White	23	31		
Ethnicity			χ^2=0 (1)	>.99
Not Latinx	23	31		
Education, n			χ^2=2.5 (3)	.47
High school diploma	1	0		
Some college	5	4		
Four-year degree	9	12		
Postgraduate degree	8	15		
PAL[c] score, mean (SD)	23.4 (8.4), 66th percentile	19.2 (8.5), 54th percentile	*W*=254 (—)	.07
median svRT[d] (ms), mean (SD)	389.7 (75), 71st percentile	418 (79), 65th percentile	*W*=451 (—)	.10
Hour of day played, mean (SD)[e]	13.1 (4.6)	14.9 (4.1)	*W*=441 (—)	.14
Hour of day played (cosine), mean (SD)	0.52 (0.6)	.45 (6)	*W*=394 (—)	.49
Super G time in target (ms), mean (SD)	653.9 (436.6)	469.4 (287.9)	*W*=279 (—)	.18
Super G response time (ms), mean (SD)	1614.8 (305.9)	1874.5 (307.4)	*W*=502 (—)	.01[b]
Super G time of reversal (ms), mean (SD)	2000.6 (592.4)	2060.3 (836.4)	*W*=432 (—)	.19
Super G scaling ratio, mean (SD)	0.66 (0.3)	0.64 (0.3)	*W*=343 (—)	.82

[a]N/A: not applicable.

[b]Statistical significance (*P*<.05).

[c]PAL: paired associates learning.

[d]svRT: simple visual reaction time.

[e]Based on 24-hour clock.

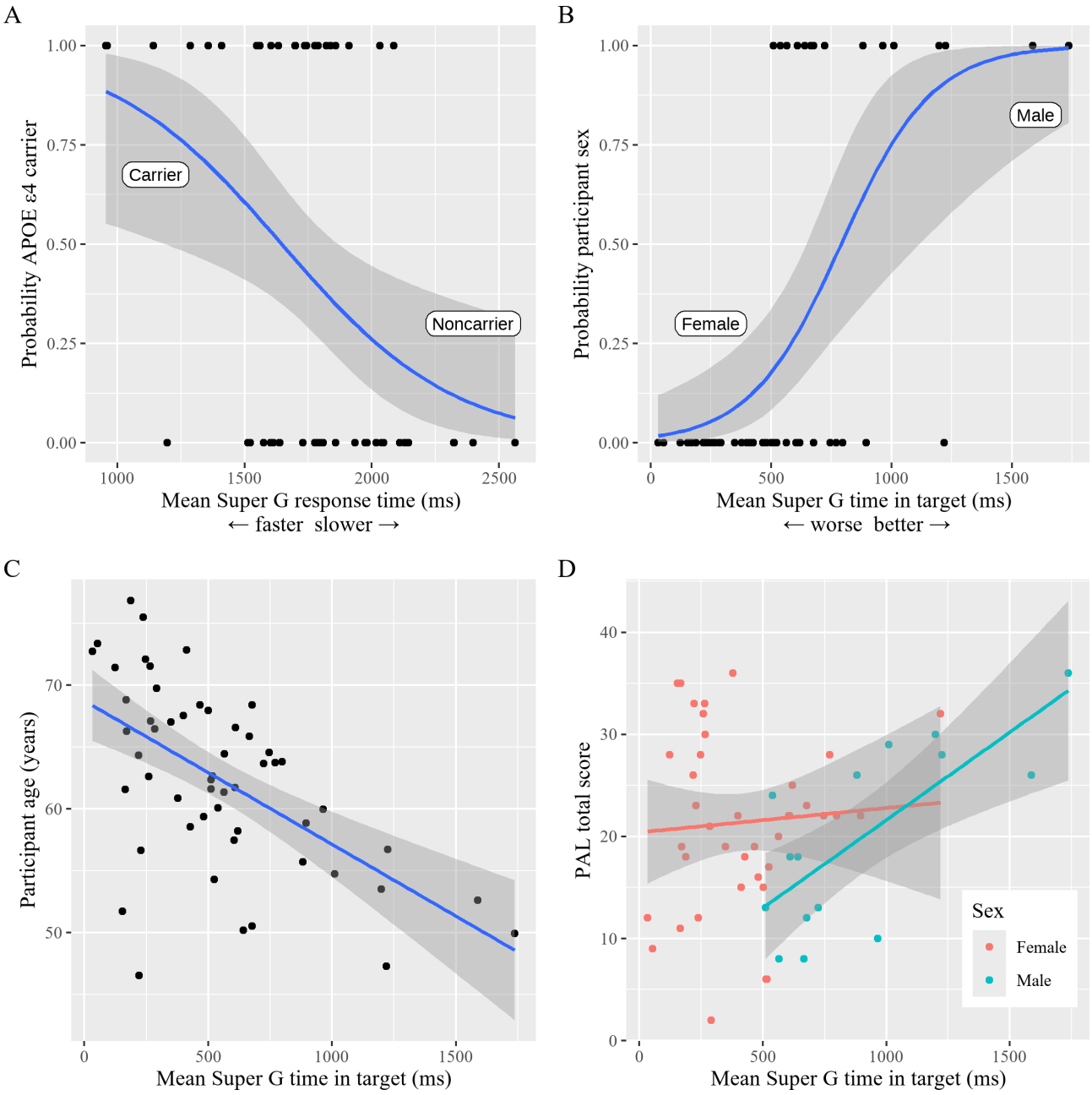
Relationship Between Super G Performance and AD Risk Factors

Within the *APOE* ε4 carrier status logistic regression (model 1) there was a significant association between *APOE* ε4 carrier status with Super G RT (OR 0.12, 95% CI 0.02-0.44; *P*=.006) (Figure 2A). Thus, a participant with a Super G RT that was 1 SD below the mean would have an 88% increase in odds of being an *APOE* ε4 noncarrier than carrier. Visual comparison between all Super G performance metrics and *APOE* ε4 can be visualized in Figure S3 in Multimedia Appendix 1. All other control variables, including PAL (*P*=.62) and svRT (*P*=.29),

were not significantly related to *APOE* ε4 carrier status. Within the participant age linear regression (model 3), Super G TinT was significantly associated with age (β −3.97, 95% CI −6.64 to −1.30; *P*=.005) (Figure 2C) and within the participant sex logistic regression (model 2), TinT was also associated with sex (OR 32.03, 95% CI 3.74-1192.61; *P*=.01) (Figure 2B), with lower TinT values (poorer performance) associated with being older and being female, respectively. This is consistent with our earlier work in other cohorts [27]. Visual comparison between all Super G performance metrics and sex can be visualized in Figure S4 in Multimedia Appendix 1.

Given that previous research in MindCrowd has demonstrated a main effect of sex on verbal learning (as measured with the PAL test) [26,28], and that TinT is also strongly linked with participant sex (Figure 2B), within the PAL linear regression (model 4), a sex-by-TinT interaction was also included to best model PAL [26]. Results from the PAL model demonstrated that age was associated with PAL (β −3.37, 95% CI −6.34 to −0.4; *P*=.03), indicating that older age was associated with lower PAL scores. There was also a main effect of participant sex (β_{Male} −7.73, 95% CI −14.98 to −0.45; *P*=.04). This result indicates that male participants scored an average of 7 points lower on the PAL than females. There was a significant sex-by-TinT interaction ($\beta_{Male:TinT}$ 6.77, 95% CI 0.34-13.19; *P*=.04). This result indicates that males with a mean TinT 1 SD greater than the group mean are associated with an increase of 6.77 points on the PAL compared with females with the same TinT performance (Figure 2D). Furthermore, in the PAL adjusted regression (model 5), where PAL scores are modified based on the time between PAL and Super G measures, the results are nearly identical to the raw scores (β_{age} −3.41, 95% CI −6.38 to −0.44; *P*=.03; β_{Male} −7.57, 95% CI −14.81 to −0.33; *P*=.04; $\beta_{Male:TinT}$ 6.72, 95% CI 0.31-13.14; *P*=.04; Figure S5 in Multimedia Appendix 1). The results of this adjusted analysis provide evidence that the delay between PAL and Super G measures does not significantly impact the observed relationship between PAL and TinT given participant sex.

Figure 2. (A) Mean Super G response time (ie, time at which Super G exits the start planet once the blue atmosphere disappears) between *APOE* ε4 noncarriers and carriers. Y-axis represents the probability a participant is classified as a carrier or noncarrier based on their mean response time. The blue line represents the binomial relationship between response time and carrier status with faster response time more associated with being a carrier and slower response time more associated with being a noncarrier. The gray ribbon represents the 95% CI of the estimated probability of binomial relationship. (B) Mean Super G time in target between female and male participants. Y-axis represents the probability a participant is classified as a male or female based on their mean time in target. The blue line represents the binomial relationship between time in target and sex with better time in target associated with being a male and worse time in target associated with being a female. The gray ribbon represents the 95% CI of the estimated probability of binomial relationship. (C) Linear relationship between participant age in years and their mean Super G time in target. The blue line represents the least squares line fitted between the variable and the gray ribbon about the line represents the 95% CI. (D)The relationship between mean Super G time in target to individual PAL total score stratified by sex (male in blue and female in pink). Given known sex differences between males and females on PAL and observed sex differences on mean Super time in target, such a stratification by sex was necessary to control for potential confounding of sex on the Super G to PAL relationship. The blue line represents the line of least squares for males and the pink line represents the line of least squares for females. The gray ribbon about each line represents the 95% CI. PAL: paired associates learning.



Discussion

This study tested whether motor skill acquisition, as assessed by the Super G task, was sensitive to known risk factors of AD, namely age, sex, *APOE* ε4 carrier status, and verbal learning. Results showed that better Super G performance was directly and independently associated with each of these risk factors, suggesting that Super G may be a sensitive digital biomarker of AD risk that can be remotely collected. Interestingly, TinT and RT were related to different AD risk factors (sex, age, and verbal learning vs *APOE* ε4 carrier status, respectively), likely reflecting how different aspects of motor skill map onto different AD risk factors. Our prior work showed that TinT is the product of optimal execution of SR and TR and best characterizes overall

skill compared with the other Super G variables [21], and it was related to age, sex, and verbal learning. However, RT is a performance characteristic of early skill acquisition, as it measures an individual's ability to anticipate when to initialize movement at the beginning of the trial, which was related to *APOE* ε4 carrier status. Although *APOE* ε4 carriers unexpectedly outperformed noncarriers in terms of early skill acquisition (ie, better RT performance), this opens up the possibility of using Super G as a prognostic enrichment strategy for enriching AD-focused cohorts with ε4 carriers [29].

The observed *APOE* ε4 benefit in this study is contrary to previous work where *APOE* ε4 carrier status leads to worse performance on memory tests in carriers compared with noncarriers [30-32]. However, the role *APOE* ε4 in aging independent of AD pathology may be significant [33], as a small but growing body of evidence in both cognitively unimpaired humans and rodents shows that visual working memory and learning is better among *APOE* ε4 carriers than noncarriers [34-36]. Although *APOE* ε4 is the primary genetic risk factor for AD [13], evidence suggests a possible benefit, or compensatory behavior [37], of learning at an earlier age while leading to impairments in later life [38,39]. Previous research has shown that cognitively unimpaired individuals who are *APOE* ε4 carriers have greater gray matter volume in frontal regions, can better allocate cognitive control, and possess better visual working memory and learning than *APOE* ε4 noncarriers [34,40,41]. Furthermore, task-based neuroimaging studies have associated better performance on visual working memory tasks among carriers with greater activation in frontal and parietal regions [34,36,42]. This leads to the hypothesis that increased frontal activation is a compensatory mechanism that modifies behavior in AD, given that frontal brain regions are relatively spared in AD [43,44]. Although these data do not provide direct support for the *APOE* ε4 compensatory mechanisms, it may be plausible that Super G may be able to measure a suspected *APOE* ε4 benefit. This interpretation is further supported by the fact that in-lab motor skill studies have associated greater skill with higher white matter integrity of frontoparietal tracts [45,46] in cognitively unimpaired older adults, which could serve as a candidate neural substrate for such a compensatory mechanism as proposed previously, but further study would be needed to confirm.

Prior research that has investigated sex differences in motor skill learning and performance has demonstrated a preferential advantage for males compared with females [47]. Similarly, this study revealed that males tended to perform better than females on Super G. The estimated effect size of mean time in target between males and females was very large, with Cohen *d*=1.66. This result is consistent with previous research using Super G, which also found large sex differences in performance [27]. Several factors may explain the observed sex differences in Super G performance. One possible biological explanation [48] may be due to early tau deposition which has been shown to be elevated in females compared with males, consistent with the higher risk for developing AD among females. Thus, a remote and unsupervised motor skill task that is sensitive to sex differences may aid in the detection of sex-specific changes in behavior due to disease in a scalable way. Moreover, there was an interaction between sex, verbal learning, and Super G performance, which suggests that motor skill in males (ie, Super G performance) may be linked to their verbal learning to a greater degree than in females. Larger sample sizes are needed to examine the interactions between sex and other behavioral variables in the context of AD [49,50].

Several limitations to this study should be acknowledged. First, the mean time between when the PAL and Super G was 5 years. Although this is a substantial delay between 2 variables of interest, there was no correlation between PAL score and this intertest interval. Furthermore, when PAL was adjusted for this delay, there was no change in the reported relationship between Super G and PAL regardless of whether raw or adjusted PAL scores were used as the dependent variable (Figure S5 in Multimedia Appendix 1). Thus, the reported result between PAL and Super G appears robust, even with such a delay between measures. Second, the study sample was all non-Hispanic White and highly educated (>80% with at least a college degree), preventing any analysis of interactions between race or ethnicity and *APOE* ε4. Analyses that consider the interaction between race and carrier status are important, given that the link between the *APOE* ε4 allele and the AD is weaker among Black/African American and Hispanic/Latino individuals despite their increased risk of developing AD overall [12,51,52]. In addition, the study sample size was relatively limited due to our required inclusion criteria, particularly the existence of *APOE* genotyping data. As such, future Super G research will recruit larger and more diverse samples. We plan to expand the size of this work with future collaborations, and it is important to note that the MindCrowd cohort itself has increased in size, in racial and ethnic diversity, and in the number of individuals in the cohort who have *APOE* genotyping data, suggesting a path to addressing this limitation in the future. Third, we did not collect participant data on previous video game experience, which may be a contributing factor to the observed relationship between Super G performance and sex [53]. Fourth, there is no defined minimal clinically important difference or clinical cutoff score for the PAL, making it difficult to determine the meaningfulness of the observed relationship between PAL and Super G performance. For example, it cannot be determined whether the predicted change in PAL of 6.5 points among males, in relation to TinT performance, is representative of a meaningful increase or decrease in verbal learning. Finally, with only an 8.1% response rate in this study (54 participants completed all trials of Super G out of the 662 contacted), there may be limited generalizability to the broader MindCrowd cohort or the general public. However, this response rate is similar to other web-based AD-focused cohorts such as the Alzheimer's Prevention Trials Webstudy [54], although the participants in this study were not paid to participate (in contrast to the Alzheimer's Prevention Trials Webstudy), were not actively seeking care, and were emailed only once during the recruitment process. It is noted that these are the factors that could influence participation rates. Despite these limitations, this study establishes the proof of concept that Super G may be a feasible skill-based digital biomarker of individual AD risk.

Acknowledgments

The authors would like to thank the participants of this study without whom this research would not be possible. This work is supported by the National Institute on Aging F32 AG071110-01A1 (principal investigator [PI]: AH), Mueller Family Charitable Trust (PI: MJH), Arizona DHS in support of the Arizona Alzheimer's Consortium (PI: Reiman), Flinn Foundation (PI: MJH), and the National Institute on Aging U19AG065169 (PI: Barnes).

Data Availability

The data, analysis, and visualizations for this paper can be found at Open Science Framework [55].

Authors' Contributions

AH participated in the original draft (lead), formal analysis (lead), writing—review and editing (equal), software (lead), methodology (equal), conceptualization (equal), and data management (equal). MJH participated in writing—review and editing (equal), conceptualization (equal), data management (equal), and methodology (equal). MDB participated in data management (equal), software (supporting), and writing—review and editing (equal). LR contributed to writing—review and editing (equal). KD contributed to writing—review and editing (equal). SYS participated in writing—review and editing (equal), conceptualization (equal), data management (equal), and methodology (equal).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental figures for between-group comparison across covariates and outcome measures contrasted with different Super G outcomes.
[PDF File, 670 KB - aging_v8i1e67298_app1.pdf]

References

1. Matthews KA, Xu W, Gaglioti AH, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. Alzheimers Dement 2019 Jan;15(1):17-24. [doi: 10.1016/j.jalz.2018.06.3063] [Medline: 30243772]
2. Klunk WE. Amyloid imaging as a biomarker for cerebral β-amyloidosis and risk prediction for Alzheimer dementia. Neurobiol Aging 2011 Dec;32(Suppl 1):S20-S36. [doi: 10.1016/j.neurobiolaging.2011.09.006] [Medline: 22078170]
3. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-β biomarkers for Alzheimer's disease. Nature New Biol 2018 Feb 8;554(7691):249-254. [doi: 10.1038/nature25456] [Medline: 29420472]
4. Chong MS, Sahadevan S. Preclinical Alzheimer's disease: diagnosis and prediction of progression. Lancet Neurol 2005 Sep;4(9):576-579. [doi: 10.1016/S1474-4422(05)70168-X] [Medline: 16109364]
5. Dagum P. Digital biomarkers of cognitive function. NPJ Digit Med 2018;1:10. [doi: 10.1038/s41746-018-0018-4] [Medline: 31304295]
6. Omberg L, Chaibub Neto E, Perumal TM, et al. Remote smartphone monitoring of Parkinson's disease and individual response to therapy. Nat Biotechnol 2022 Apr;40(4):480-487. [doi: 10.1038/s41587-021-00974-9] [Medline: 34373643]
7. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. Ann Neurol 2024 Mar;95(3):507-517. [doi: 10.1002/ana.26833] [Medline: 37991080]
8. Mueller KD, Van Hulle CA, Koscik RL, et al. Amyloid beta associations with connected speech in cognitively unimpaired adults. Alzheimers Dement (Amst) 2021;13(1):e12203. [doi: 10.1002/dad2.12203] [Medline: 34095435]
9. Rentz DM, Papp KV, Mayblyum DV, et al. Association of digital clock drawing with PET amyloid and tau pathology in normal older adults. Neurology (ECronicon) 2021 Apr 6;96(14):e1844-e1854. [doi: 10.1212/WNL.0000000000011697] [Medline: 33589537]
10. Gurland BJ, Wilder DE, Lantigua R, et al. Rates of dementia in three ethnoracial groups. Int J Geriatr Psychiatry 1999 Jun;14(6):481-493. [Medline: 10398359]
11. Podcasy JL, Epperson CN. Considering sex and gender in Alzheimer disease and other dementias. Dialogues Clin Neurosci 2016 Dec;18(4):437-446. [doi: 10.31887/DCNS.2016.18.4/cepperson] [Medline: 28179815]
12. Tang MX, Stern Y, Marder K, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. JAMA 1998 Mar 11;279(10):751-755. [doi: 10.1001/jama.279.10.751] [Medline: 9508150]
13. Liu CC, Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol 2013 Feb;9(2):106-118. [doi: 10.1038/nrneurol.2012.263] [Medline: 23296339]
14. Schmidt R, Lee T, Winstein C, Wulf G, Zelaznik H. Motor Control and Learning: A Behavioral Emphasis, 6th edition: Human Kinetics; 2019.

15. Rogojin A, Gorbet DJ, Hawkins KM, Sergio LE. Cognitive-motor integration performance is affected by sex, APOE status, and family history of dementia. J Alzheimers Dis 2019;71(2):685-701. [doi: 10.3233/JAD-190403]
16. Schaefer SY, Malek-Ahmadi M, Hooyman A, King JB, Duff K. Association between motor task performance and hippocampal atrophy across cognitively unimpaired, amnestic mild cognitive impairment, and Alzheimer's disease individuals. J Alzheimers Dis 2022;85(4):1411-1417. [doi: 10.3233/JAD-210665] [Medline: 34958015]
17. Schaefer SY, Hooyman A, Duff K. Using a timed motor task to predict one-year functional decline in amnestic mild cognitive impairment. J Alzheimers Dis 2020;77(1):53-58. [doi: 10.3233/JAD-200518] [Medline: 32651327]
18. Schaefer SY, Duff K, Hooyman A, Hoffman JM. Improving prediction of amyloid deposition in mild cognitive impairment with a timed motor task. Am J Alzheimers Dis Other Demen 2022;37:15333175211048262. [doi: 10.1177/15333175211048262] [Medline: 35200059]
19. Tsay JS, Lee A, Ivry RB, Avraham G. Moving outside the lab: the viability of conducting sensorimotor learning studies online. Neuron Behav Data Anal Theory 2021;5(3). [doi: 10.51628/001c.26985]
20. Hooyman A, Huentelman MJ, De Both M, Ryan L, Schaefer SY. Establishing the validity and reliability of an online motor learning game: applications for Alzheimer's disease research within MindCrowd. Games Health J 2023 Apr;12(2):132-139. [doi: 10.1089/g4h.2022.0042] [Medline: 36745382]
21. Hooyman A, Gordon J, Winstein C. Unique behavioral strategies in visuomotor learning: hope for the non-learner. Hum Mov Sci 2021 Oct;79:102858. [doi: 10.1016/j.humov.2021.102858] [Medline: 34392189]
22. Talboom JS, De Both MD, Naymik MA, et al. Two separate, large cohorts reveal potential modifiers of age-associated variation in visual reaction time performance. NPJ Aging Mech Dis 2021 Jul 1;7(1):14. [doi: 10.1038/s41514-021-00067-6] [Medline: 34210964]
23. Lewis CR, Talboom JS, De Both MD, et al. Smoking is associated with impaired verbal learning and memory performance in women more than men. Sci Rep 2021 May 13;11(1):10248. [doi: 10.1038/s41598-021-88923-z] [Medline: 33986309]
24. Brooks V, Hilperath F, Brooks M, Ross HG, Freund HJ. Learning "what" and "how" in a human motor task. Learn Mem 1995;2(5):225-242. [doi: 10.1101/lm.2.5.225]
25. Pike KE, Rowe CC, Moss SA, Savage G. Memory profiling with paired associate learning in Alzheimer's disease, mild cognitive impairment, and healthy aging. Neuropsychology 2008 Nov;22(6):718-728. [doi: 10.1037/a0013050] [Medline: 18999345]
26. Talboom JS, Håberg A, De Both MD, et al. Family history of Alzheimer's disease alters cognition and is modified by medical and genetic factors. Elife 2019 Jun 18;8:e46179. [doi: 10.7554/eLife.46179] [Medline: 31210642]
27. Hooyman A, Schaefer SY. Age and sex effects on Super G performance are consistent across internet devices. Int J Serious Games 2023 Jun 7;10(2):25-36. [doi: 10.17083/ijsg.v10i2.598] [Medline: 37846217]
28. Kaushanskaya M, Marian V, Yoo J. Gender differences in adult word learning. Acta Psychol (Amst) 2011 May;137(1):24-35. [doi: 10.1016/j.actpsy.2011.02.002] [Medline: 21392726]
29. Cummings J, Apostolova L, Rabinovici GD, et al. Lecanemab: appropriate use recommendations. J Prev Alzheimers Dis 2023;10(3):362-377. [doi: 10.14283/jpad.2023.30] [Medline: 37357276]
30. Caselli RJ, Reiman EM, Osborne D, et al. Longitudinal changes in cognition and behavior in asymptomatic carriers of the APOE e4 allele. Neurology (ECronicon) 2004 Jun 8;62(11):1990-1995. [doi: 10.1212/01.wnl.0000129533.26544.bf] [Medline: 15184602]
31. Marioni RE, Campbell A, Scotland G, Hayward C, Porteous DJ, Deary IJ. Differential effects of the APOE e4 allele on different domains of cognitive ability across the life-course. Eur J Hum Genet 2016 Jun;24(6):919-923. [doi: 10.1038/ejhg.2015.210] [Medline: 26395552]
32. Evans SL, Dowell NG, Prowse F, Tabet N, King SL, Rusted JM. Mid age APOE ε4 carriers show memory-related functional differences and disrupted structure-function relationships in hippocampal regions. Sci Rep 2020 Feb 20;10(1):3110. [doi: 10.1038/s41598-020-59272-0] [Medline: 32080211]
33. Palmer JM, Huentelman M, Ryan L. More than just risk for Alzheimer's disease: APOE ε4's impact on the aging brain. Trends Neurosci 2023 Sep;46(9):750-763. [doi: 10.1016/j.tins.2023.06.003] [Medline: 37460334]
34. Lu K, Nicholas JM, Pertzov Y, et al. Dissociable effects of APOE-ε4 and β-amyloid pathology on visual working memory. Nat Aging 2021 Nov;1(11):1002-1009. [doi: 10.1038/s43587-021-00117-4] [Medline: 34806027]
35. McLean JW, Bhattrai A, Vitali F, Raikes AC, Wiegand JP, Brinton RD. Contributions of sex and genotype to exploratory behavior differences in an aged humanized APOE mouse model of late-onset Alzheimer's disease. Learn Mem 2022 Sep;29(9):321-331. [doi: 10.1101/lm.053588.122]
36. Scheller E, Peter J, Schumacher LV, et al. APOE moderates compensatory recruitment of neuronal resources during working memory processing in healthy older adults. Neurobiol Aging 2017 Aug;56:127-137. [doi: 10.1016/j.neurobiolaging.2017.04.015] [Medline: 28528773]
37. Filbey FM, Chen G, Sunderland T, Cohen RM. Failing compensatory mechanisms during working memory in older apolipoprotein E-epsilon4 healthy adults. Brain Imaging Behav 2010 Jun;4(2):177-188. [doi: 10.1007/s11682-010-9097-9] [Medline: 20502990]
38. Iacono D, Feltis GC. Impact of apolipoprotein E gene polymorphism during normal and pathological conditions of the brain across the lifespan. Aging (Albany NY) 2019 Jan 24;11(2):787-816. [doi: 10.18632/aging.101757] [Medline: 30677746]

39. Palmer JM, Grilli MD, Lawrence AV, Ryan L. The impact of context on pattern separation for objects among younger and older apolipoprotein 4 carriers and noncarriers. J Int Neuropsychol Soc 2023 Jun;29(5):439-449. [doi: 10.1017/S1355617722000455] [Medline: 36416211]
40. Dean DC III, Jerskey BA, Chen K, et al. Brain differences in infants at differential genetic risk for late-onset Alzheimer disease. JAMA Neurol 2014 Jan 1;71(1):11. [doi: 10.1001/jamaneurol.2013.4544]
41. Zink N, Bensmann W, Arning L, Beste C, Stock AK. Apolipoprotein ε4 is associated with better cognitive control allocation in healthy young adults. Neuroimage 2019 Jan 15;185:274-285. [doi: 10.1016/j.neuroimage.2018.10.046] [Medline: 30342978]
42. Wishart HA, Saykin AJ, Rabin LA, et al. Increased brain activation during working memory in cognitively intact adults with the APOE epsilon4 allele. Am J Psychiatry 2006 Sep;163(9):1603-1610. [doi: 10.1176/ajp.2006.163.9.1603] [Medline: 16946187]
43. Tuminello ER, Han SD. The apolipoprotein E antagonistic pleiotropy hypothesis: review and recommendations. Int J Alzheimers Dis 2011 Feb 24;2011:726197. [doi: 10.4061/2011/726197] [Medline: 21423560]
44. Han SD, Bondi MW. Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. Alzheimers Dement 2008 Jul;4(4):251-254. [doi: 10.1016/j.jalz.2008.02.006] [Medline: 18631975]
45. Lingo VanGilder J, Bergamino M, Hooyman A, et al. Using whole-brain diffusion tensor analysis to evaluate white matter structural correlates of delayed visuospatial memory and one-week motor skill retention in nondemented older adults: a preliminary study. PLoS One 2022;17(9):e0274955. [doi: 10.1371/journal.pone.0274955] [Medline: 36137126]
46. Regan E, Fridriksson J, Schaefer SY, et al. Neural correlates of within-session practice effects in mild motor impairment after stroke: a preliminary investigation. Exp Brain Res 2021 Jan;239(1):151-160. [doi: 10.1007/s00221-020-05964-y] [Medline: 33130906]
47. Moreno-Briseño P, Díaz R, Campos-Romo A, Fernandez-Ruiz J. Sex-related differences in motor learning and performance. Behav Brain Funct 2010 Dec 23;6(1):74. [doi: 10.1186/1744-9081-6-74] [Medline: 21182785]
48. Buckley RF, Mormino EC, Rabin JS, et al. Sex differences in the association of global amyloid and regional tau deposition measured by positron emission tomography in clinically normal older adults. JAMA Neurol 2019 May 1;76(5):542-551. [doi: 10.1001/jamaneurol.2018.4693] [Medline: 30715078]
49. Caselli RJ, Langlais BT, Dueck AC, et al. Neuropsychological decline up to 20 years before incident mild cognitive impairment. Alzheimers Dement 2020 Mar;16(3):512-523. [doi: 10.1016/j.jalz.2019.09.085] [Medline: 31787561]
50. Savignac C, Villeneuve S, Badhwar A, et al. APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementia. PLoS Biol 2022 Dec;20(12):e3001863. [doi: 10.1371/journal.pbio.3001863] [Medline: 36512526]
51. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. JAMA 1997;278(16):1349-1356. [Medline: 9343467]
52. Harwood DG, Ownby RL. Ethnicity and dementia. Curr Psychiatry Rep 2000 Feb;2(1):40-45. [doi: 10.1007/s11920-000-0040-4] [Medline: 11122930]
53. Terlecki M, Brown J, Harner-Steciw L, et al. Sex differences and similarities in video game experience, preferences, and self-efficacy: implications for the gaming industry. Curr Psychol 2011 Mar;30(1):22-33. [doi: 10.1007/s12144-010-9095-5]
54. Walter S, Clanton TB, Langford OG, et al. Recruitment into the Alzheimer Prevention Trials (APT) Webstudy for a Trial-Ready Cohort for Preclinical and Prodromal Alzheimer's Disease (TRC-PAD). J Prev Alzheimers Dis 2020;7(4):219-225. [doi: 10.14283/jpad.2020.46] [Medline: 32920623]
55. Hooyman A. Super G and AD risk factors in Mindcrowd. OSF. URL: https://osf.io/rnzg6 [accessed 2025-03-13]

Abbreviations

AD: Alzheimer disease
MCI: mild cognitive impairment
OR: odds ratio
PAL: paired associates learning
RT: response time
SR: scaling ratio
svRT: simple visual reaction time
TinT: time in the target
TR: time of reversal

Edited by A Garcia; submitted 08.10.24; peer-reviewed by A Lewis, AC Cavalcanti Maciel; revised version received 24.02.25; accepted 11.03.25; published 24.04.25.

Please cite as:
Hooyman A, Huentelman MJ, De Both M, Ryan L, Duff K, Schaefer SY
Relationship Between Within-Session Digital Motor Skill Acquisition and Alzheimer Disease Risk Factors Among the MindCrowd Cohort: Cross-Sectional Descriptive Study
JMIR Aging 2025;8:e67298
URL: https://aging.jmir.org/2025/1/e67298
doi:10.2196/67298

Original Paper

Implications of Public Disclosure of Personal Information in a Mobile Alert App for People Living With Dementia Who Go Missing: Qualitative Descriptive Study

Adebusola Adekoya[1], MN; Christine Daum[1], PhD; Noelannah Neubauer[1], PhD; Antonio Miguel-Cruz[2], PhD; Lili Liu[1], PhD

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

[2]Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Lili Liu, PhD
School of Public Health Sciences, Faculty of Health
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567
Email: lili.liu@uwaterloo.ca

Abstract

Background: People living with dementia are at risk of getting lost and going missing due to memory loss, confusion, and disorientation. Missing person incidents involving people living with dementia are increasing. Alert systems such as Community ASAP can promote community engagement in locating missing persons with dementia and aid in search and rescue efforts. However, the implications of public disclosure of personal information such as name, age, sex, and physical description within such alert systems have yet to be explored.

Objective: This study aimed to identify and discuss the implications of public disclosure of personal information in Community ASAP for people living with dementia at risk of going missing.

Methods: This study used a qualitative descriptive research design drawing from naturalistic inquiry. A total of 19 participants including people living with dementia, care partners, first responders, and service providers were recruited from Ontario, Alberta, and British Columbia, Canada. Semistructured interviews were used to explore participants' perspectives on the perceived implications of the release of personal information when using Community ASAP. NVivo (version 12) was used to manage data, and conventional content analysis was conducted to identify key themes of the implications of public disclosure of personal information in Community ASAP.

Results: In total, 10/19 (53%) of the participants were women and 9/19 (47%) were men. Of the 19 participants, 3 (16%) were people living with dementia, 5 (26%) were care partners, 4 (21%) were first responders, and 7 (37%) were service providers. In total, 4 key themes were identified as implications of public disclosure of personal information in Community ASAP: *right to autonomy*, *safety versus privacy*, *informed and knowledgeable consent*, and *stigmatization.* Participants discussed how the public disclosure of personal information in Community ASAP could undermine a person's choice not to be found and contribute to stigmatization. Participants emphasized a need to balance safety and privacy concerns. Informed and knowledgeable consent is important when using an alert system to locate missing persons with dementia.

Conclusions: Community ASAP can promote community engagement in locating missing persons with dementia. However, the public disclosure of personal information in alert systems has implications. Users' right to autonomy, a balance between safety and privacy, informed and knowledgeable consent, and risks of stigmatization are perceived impacts of disclosure of personal information in alert systems.

(JMIR Aging 2025;8:e64847) doi:10.2196/64847

KEYWORDS

alert systems; technology; missing persons; dementia; autonomy; privacy; stigmatization; consent

Introduction

Background

Dementia is an increasing health problem affecting >55 million people worldwide [1]. People living with dementia are at risk of getting lost and going missing due to memory loss, confusion, and disorientation [2]. Missing person incidents involving people living with dementia have been increasing [3]. Missing persons with dementia are at risk of being exposed to injuries and death. If not found within 24 hours, half of missing persons with dementia will sustain serious injuries or be found deceased [4]. It is crucial to engage the community in locating missing persons and not view missing person incidents solely as a family issue [5]. This inclusive approach provides extra eyes on the ground. Strategies such as alert systems promote community engagement by encouraging community members to be on the lookout for missing persons. This, in turn, aids in search and rescue efforts and minimizes the risks associated with going missing [6].

Alert systems such as the United States' publicly funded Silver Alert program disseminate information about missing persons with dementia or other cognitive disabilities to the public through media broadcasts (eg, commercial radio and television stations) and electronic billboards [6]. The Silver Alert system has been implemented in all but 5 states [6]. However, there are limitations to these programs. The process of issuing Silver Alerts and the duration differ across states, and media broadcasts may reach fewer people due to an increase in the use of mobile devices to access information. Furthermore, there are concerns about declining media sensitivity due to alert fatigue [7].

In Canada, when a person living with dementia goes missing, police services alert the public about a missing person using social media such as Facebook and X (formerly Twitter) and not an emergency alert system such as an Amber alert [8]. Amber alerts notify the public about missing abducted children using Alert Ready. Alert Ready is Canada's emergency alert system that sends critical alerts about hazards or impending dangers (eg, floods, tornados, chemical spills, and fires) to Canadians through radio, satellite television, and compatible wireless devices [9]. As missing person incidents involving people living with dementia are more frequent than missing children and not typically related to a crime, the use of the Alert Ready system would not be appropriate [8]. Furthermore, the use of the Alert Ready system can desensitize the public to alerts and cause alert fatigue, which already exists with Amber alerts.

Ideally, programs that mitigate risks of going missing among people living with dementia would engage local communities but minimize alert fatigue and be location specific. For example, the community-led British Columbia Silver Alert program notifies community subscribers about missing persons with dementia, cognitive impairment, and autism using channels such as social media (eg, Facebook and X) and via email or SMS text message [10]. Interest in community alert systems in other Canadian provinces is growing, as evidenced by a recent national petition; amendments to Missing Persons (Silver Alert) Acts in Alberta, Manitoba, and Ontario [11-13]; and the launch of Silver Alert pilot projects in Quebec [14].

To date, a localized alert system based on volunteer subscription does not exist. To this end, our team developed a localized area alert app called Community ASAP, the first of its kind in Canada, with representatives of end users, including people living with dementia, care partners, and community organizations [3]. Community ASAP allows community volunteers, local businesses, and police services to work together to locate missing persons with cognitive impairment and is available as a mobile app. As a requirement in Canada, a Community ASAP coordinator from police services would initiate an alert when a person living with dementia is reported missing. Community volunteers who register with Community ASAP receive alerts about missing persons based on their geographic preferences. The alert includes a link to the missing individual's personal information, such as their first and last names, nickname, age, physical description (eg, height, weight, eye color, and hair color), mobile number (if any), and locations they typically visit [3].

Community ASAP and other public alert systems such as Silver Alerts and Amber alerts are designed to enhance public awareness of vulnerable missing persons, but they differ in their scope and approaches. Community ASAP is a subscription-based alert system that requires community engagement and provides real-time updates and geofencing features. For example, volunteers receive missing person alerts based on their geographic preferences, allowing them to choose up to 5 locations, such as home and work addresses, and set a radius (1, 3, 6, 12, or 25 km) for each location [3]. However, its effectiveness depends on widespread app adoption and access to mobile devices. In contrast, Silver Alerts and Amber alerts are government funded or community funded and broadcast information about missing persons through large-scale public channels such as highway signs, radio, and television, reaching a wider audience [6,15]. The activation of Community ASAP in Canada and public alert systems in the United States rely on police verification, which means that alerts are not triggered immediately when a missing person incident occurs [3,6,16].

This study was a part of the development and evaluation of the accuracy and usability of Community ASAP [3]. During the development of Community ASAP, we identified the need to understand how the rights of people living with dementia to self-determination and privacy are respected while using the alert system. The use of alert systems has raised ethical concerns about the privacy of people living with dementia [6,15]. Furthermore, there is limited knowledge about concerns related to using alert systems to locate missing persons with dementia [15], particularly how the release of personal information to the public may impinge on the human rights of vulnerable older adults [17].

Objectives

The purpose of this study was to identify and discuss the implications of public disclosure of personal information in an alert system called Community ASAP for people living with dementia at risk of going missing.

Methods

Research Design

This study used a qualitative descriptive research design drawing from naturalistic inquiry, which allows a phenomenon to be studied in its natural state [18]. This method is appropriate when the intent of the study is to provide a basic description and summary of people's experiences [19,20].

Recruitment

Purposeful sampling [21] was used to select participants who had professional and personal experiences on the topic to best provide insights into concerns that could arise from the release of information via Community ASAP. Specific strategies included intensity [21] and snowball [22] sampling. Intensity sampling was used to sample information-rich cases. In snowball sampling, participants helped share study information with other prospective participants whom they believed had knowledge or experience with the topic of interest. These strategies were chosen for practical reasons, including the challenges of selecting potentially difficult-to-access or vulnerable groups, such as people living with dementia. We aimed to include the perspectives of people living with dementia, who are often underrepresented in research. Studies have shown the importance of involving people living with dementia in research, highlighting their ability to provide reliable self-report information, personal perspectives, and valuable insights [23-25]. Inclusion of people living with dementia in this study was crucial to understanding their individual experiences with an alert system and enhancing their safety and well-being while using this program.

Participants (people living with dementia, care partners, first responders, and service providers) were recruited and identified via email or face-to-face conversations through our research team's existing professional networks serving people living with dementia, such as Alzheimer societies, dementia advocacy organizations, and first responders. Thus, some of the participants were known to the research team, and others were not. None of the participants were family members or friends of or had personal relationships with the researchers. All prospective participants who were invited to take part in the study enrolled. Participants had lived or professional experience with dementia and the use of technologies to manage the risk of going missing or responding to missing person incidents. During recruitment, participants were asked to self-identify their experience and knowledge of technologies. All participants had some level of familiarity or experience with technology. We did not inquire about the extent or specific type of their experience.

We selected participants from across 3 Canadian provinces (Ontario, Alberta, and British Columbia) where our professional networks were located and from 4 stakeholder groups: people living with dementia, care partners, service providers (eg, social workers, dementia educators, representatives from vulnerable person registries, and support workers), and first responders (eg, police and search and rescue managers). Inclusion criteria were to (1) be aged ≥18 years, (2) speak English, and (3) have lived or professional experience with or knowledge of the use of technologies to manage the risk of going missing or responding to missing person incidents. Exclusion criteria were having no understanding about technologies to manage the risk of going missing, inability to articulate perspectives due to impaired cognitive abilities, and severe visual or hearing limitations that could not be corrected with the use of an assistive device.

Ethical Considerations

Ethics approval was received from the University of Alberta Research Ethics Board (Pro00078537). All participants gave written informed consent prior to the study, were told that their participation is voluntary, and were informed of their right to withdraw at any time. People living with mild cognitive impairment or mild dementia who had the ability to consent as identified by our professional networks were invited to participate in this study. We used the teach-back method [26] to ensure that these participants had cognitive ability to engage in one-on-one interviews with research team members. The person living with dementia was asked an open-ended question about what they had read in the information letter to ensure that they understood the procedure, risks, and what to do should they wish to withdraw from the study. Transcripts were cleaned for accuracy and deidentified to maintain anonymity. Participants did not receive an honorarium.

Data Collection

Semistructured interviews [27] were conducted with each participant in person, via phone, or through Zoom (Zoom Video Communications) videoconference. The purpose of the interviews was to understand participant perspectives on the implications of public disclosure of personal information in alert systems. Each person selected a preferred mode of participation based on their geographical location and other time commitments. In total, we interviewed 19 stakeholders (n=16, 84% individual interviews and 1 group interview with n=3, 16% of the participants at their request). No other individuals (ie, care recipients or care partners) were present during the interviews with people living with dementia as these individuals had the ability to consent and respond to questions. Before the interview began, the researcher introduced herself, the overarching study [3] in which the current project was embedded, and the purpose and procedures, and written informed consent was obtained. An interview guide (Textbox 1) that contained 5 open-ended questions was used to elicit participants' perspectives on the possible implications of the release of information through Community ASAP. Probes were used to elaborate on participant responses and clarify meaning [22]. The interviews were conducted by female research team members (NN and CD) with backgrounds in occupational therapy, experience with qualitative research methods, and PhDs in rehabilitation science. Each participant was interviewed once, and the interviews were approximately 30 to 90 minutes in length. Interviewers documented their observations and any interruptions in field notes, and the interviews were digitally recorded and transcribed verbatim. Transcripts were not returned to participants as we did not use a member-checking approach to review the transcripts. NVivo (version 12; Lumivero) was used to manage the data.

Textbox 1. Interview guide questions.

Questions

- From your perspective as a (person with dementia or caregiver of a person with dementia, service provider, first responder, expert in ethics, or the law), what ethical concerns are associated with release of personal information in the C-ASAP system?
- From your perspective as a (person with dementia or caregiver of a person with dementia, service provider, first responder, expert in ethics, or the law), what legal concerns are associated with release of personal information in the C-ASAP system?
- How does the release of personal information in the proposed C-ASAP system compare with other registries such as MedicAlert Safely Home and various Vulnerable Persons Registries?
- What are the implications of having a missing person's name released to the public (and thus a part of the public record)? What is the balance (or tipping point) between privacy and safety?
- [For group interviews involving representatives from agencies that collect information about vulnerable persons and store them on registries]: What agreements are in place between the vulnerable person or their family and the agency upon signing on with a particular registry?

Data Analysis

Conventional content analysis [28] guided our analytic approach. A thematic approach is preferred when there is limited existing theory or research literature on a phenomenon, allowing for the identification of specific meanings and determination of appropriate categories and themes. A female doctoral candidate in public health sciences (AA) with a background in nursing and qualitative research methods analyzed the data. She immersed herself in the data by listening to the recordings while reading the transcripts and identifying initial reflections. Next, she coded the transcripts using keywords. Codes were described, and similar codes were grouped and refined to create categories. As new codes were generated, they were added to the framework, and transcripts that had been previously coded were updated to reflect hierarchy. Category descriptions were generated to describe their contents, and themes that represented the categories were created inductively. Analysis continued until saturation of the data was reached.

Data analysis was an iterative process. The coding hierarchy, categories, and themes were reviewed, scrutinized, and confirmed by the analyst and the 2 team members who conducted the interviews; that is, we used peer debriefing as a trustworthiness strategy [29]. This helped us improve clarity and the internal and external homogeneity of codes and categories [30] and consider alternative interpretations and explanations. Thus, the coding hierarchy and main themes were refined repeatedly, thereby enhancing credibility. Our process is consistent with the approach by Morse et al [31] to verification in which transcripts, codes, categories, and themes are rechecked.

Results

Overview

A total of 19 participants (n=10, 53% women and n=9, 47% men) were interviewed. Of these 19 participants, 3 (16%) were people living with dementia, 5 (26%) were care partners, 4 (21%) were first responders (search and rescue and police officers), and 7 (37%) were service providers. Participants were predominantly White individuals (17/19, 89%), and the remainder were Asian individuals (Korean and Filipino; 2/19, 11%; Table 1). All participants had some level of familiarity or experience with technology. In total, 58% (11/19) of the participants (first responders and service providers) had professional experience with technology to manage or respond to missing person incidents, such as alert systems, MedicAlert, and locator devices (GPS devices and Project Lifesaver). The remaining participants (people living with dementia and care partners; 8/19, 42%) had lived experience or familiarity these technologies.

Table 1. Study participants.

	Care partners (n=5), n (%)	First responders (n=4), n (%)	People living with dementia (n=3), n (%)	Service providers (n=7), n (%)
Sex				
Male	4 (80)	4 (100)	1 (33)	0 (0)
Female	1 (20)	0 (0)	2 (67)	7 (100)
Ethnicity				
Filipino	0 (0)	0 (0)	0 (0)	1 (14)
Korean	1 (20)	0 (0)	0 (0)	0 (0)
White	4 (80)	4 (100)	3 (100)	6 (86)
Province				
Alberta	1 (20)	1 (25)	0 (0)	0 (0)
British Columbia	1 (20)	1 (25)	1 (33)	2 (29)
Ontario	3 (60)	2 (50)	2 (67)	5 (71)

Thematic Findings

Overview

Four key themes represent the implications of public disclosure of personal information in Community ASAP for people living with dementia who have gone missing: (1) right to autonomy, (2) safety versus privacy, (3) informed and knowledgeable consent, and (4) stigmatization (Textbox 2).

Textbox 2. Themes with participant quotes.

Right to autonomy

- "Depending on where the person is in their journey, what if they just wanted to be alone for a few days and then got this whole thing going on because someone decided that they've disappeared. That's a legal issue for me." [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]
- "We locate a missing person, we let them know you've been reported missing, here's the person who's reported you missing...However, they do have a right to be missing so to speak." [First responder 4, an experienced police officer with expertise in search and rescue]
- "Is that what they wanted when they were in the real self before dementia, there is an ethical issue there because you're not honoring what the person wants in the moment." [Service provider 3, a service provider who worked directly with people living with dementia providing education and counseling]

Safety versus privacy

- "The more eyes the better. I think if you're missing, I don't think privacy should be an issue." [Person living with dementia 1, who had experience with MedicAlert]
- "I think that the tipping point is safety, and so that might be it is 30 to 20 degrees Celsius outside. It's not as extreme situation as it is 30 or minus 25. So, all of these things need to be taken into account." [Care partner 2, a family member who cared for a person living with dementia]
- "Some people aren't as open to having other people or the general public know a person with dementia has been missing more around, that they have the condition, they have dementia." [Service provider 6, a person with expertise in supporting the safety of vulnerable persons and who had experience with GPS locator devices for this population]
- "If they are making this phone call it's urgent enough, it's important enough that yes, we are willing to give up that right to this in order to find this person quickly." [First responder 1, an experienced search and rescue member who also provided education about locator devices]

Stigmatization

- "For a lot of people, I worry about their faces and the local papers, this stuff. And that still makes them very vulnerable, because someone recognizes them after the fact." [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]
- "So, with my dad, we didn't want anybody to know that he had Alzheimer's. I think it's because there may be a stigma of Alzheimer's." [Care partner 3, a family member of a person living with dementia who got lost]
- "So, it's such a stigma, it's such a huge issue that sometimes you might think that having an alert system would be really useful for the community to know how to interact or to keep their eye out for somebody with dementia in the community. But is a stigma sort of perpetuating." [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]
- "The one thing we are always cognizant of in our media releases, is our management and our upper management are always cognizant of the...stigma." [First responder 3, a police officer with experience with missing persons]

Informed and knowledgeable consent

- "If the person is like me and can give consent, that would be OK...if I've already deteriorated, then I think it would be up to my family or my power of attorney because they know what my wishes are." [Person living with dementia 2, a person who was previously lost]
- "Whenever possible, that person who is actually living with dementia, I certainly would advocate and support their involvement from the get-go. No matter what the timeframe, I think you need to be respectful and try to include their voice as much as possible, knowing that cognitively that could change at any time, and if it does then I would hope that the person has the best interests for that individual." [Care partner 5, a family member of a person living with dementia who had experience with locator devices]
- "I think the idea of capacity and the ability to give informed consent is not an on/off button. And to always have that conversation as much as possible to the extent that it is possible for you and whether that's through care partners. I think it's really at the heart of it, really enabling that person to have the choice. But the challenge with that is that on one day the person might say 'yeah I'm completely fine with it' and then the next day they're like 'ugh, I'm not fine with it at all.'" [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]
- "Informed consent is the only consent that would certainly help alleviate some problems down the road, and I would think if the person can't give consent, because of their mental capacity, hopefully the caregiver would have that Power of Attorney to be able to make that decision for them, on their behalf." [First responder 2, an experienced police officer and search and rescue expert]

Right to Autonomy

This theme relates to concerns about how the release of personal information can compromise a person's choice not to be found. Participants discussed the importance of respecting people's autonomy and right to go missing without wanting to be found. People living with dementia also possess the rights of adults and have the right to be alone or go missing, whereas this does not apply to children. A participant (first responder 4) with many years of experience in search and rescue emphasized adults' rights to go missing. He expressed the following:

> *When we locate a missing person, we let them know you've been reported missing, here's the person who's reported you missing...However, they do have a right to be missing so to speak.*

Labeling intentionally missing persons as missing, especially those capable of decision-making, may raise ethical and legal concerns, as noted by some participants. This concern was further expanded by a participant living with dementia:

> *Depending on where the person is in their journey, what if they just wanted to be alone for a few days and then got this whole thing going on because someone decided that they've disappeared. That's a legal issue for me. Like where in the journey it is decided that someone...I don't even know how to explain that because like for me, it's not uncommon for me, I mean now I always tell somebody just because that's my life. I quite often go "Oh well I'm going to take off and go here for a couple of days."* [Person living with dementia 3, an experienced user of GPS devices who had previously been lost]

Participants pointed out that individuals' views on risk and outlook on life may change as their dementia progresses. Hence, it is crucial that the use of Community ASAP respects the preferences of those living with dementia. In addition, there might be a need to re-evaluate personal information disclosure preferences as dementia progresses. A service provider highlighted the importance of honoring the current wishes of people living with dementia when using Community ASAP:

> *Somebody's idea about what they consent to and what they want when they're first diagnosed might be different than what they want later on, and what's the record of truth. Is that what they wanted when they were in the real self before dementia, there is an ethical issue there because you're not honoring what the person wants in the moment.* [Service provider 3, a service provider who worked directly with people living with dementia providing education and counseling]

Safety Versus Privacy

This theme highlights concerns related to the balance between safety and privacy, acknowledging the diverse perspectives on how to maintain this balance for missing persons. While Community ASAP serves as a safety net, offering peace of mind to people living with dementia and their care partners, it also facilitates a sense of independence. However, some participants expressed concerns about privacy when it comes to sharing their personal information publicly. For many, safety was the tipping point as finding the missing person alive and minimizing the risks of harm is paramount. A care partner emphasized the urgency of finding missing persons promptly, especially in severe weather conditions in Canada:

> *I think that the tipping point is safety, and so that might be it is 30 to 20 degrees Celsius outside. It's not as extreme situation as it is 30 or minus 25. So, all of these things need to be taken into account.* [Care partner 2, a family member who cared for a person living with dementia]

Privacy concerns are influenced by culture as cultural beliefs play a key role in how individuals view privacy, particularly regarding the disclosure of certain information to the public. For instance, a Korean participant (care partner 3, a family member of a person living with dementia who got lost) shared cultural beliefs about having a "conservative background" and keeping his father's dementia diagnosis confidential within the family and refraining from sharing it with outsiders.

There are concerns regarding the permanence of personal information on the internet. Details such as a missing person's name, age, photo, and physical description shared with the public might linger on the internet even after the person is located. Some participants expressed unease about their personal information being public posthumously.

Participants also emphasized the importance of only sharing essential information, such as the person's name, age, photo, and physical description. While some information is necessary for community members (volunteers) to identify a missing individual, disclosing medical information (eg, cognitive impairment) could heighten the risks of fraud, identity theft, and abuse. One participant raised concerns about potential exploitation of people living with dementia when using Community ASAP:

> *The risk is that the volunteers would prey upon this person if he/she is found and returned home, identified to have dementia. They are already vulnerable to scammers, to people who would be able to redirect them and take advantage of them.* [Service provider 1, a service provider who worked directly with people living with dementia providing education and counseling]

Some participants expressed that a dementia diagnosis could make individuals vulnerable to exploitation regardless of whether their personal information is shared via an alert system such as Community ASAP. One participant (person living with dementia 1) stressed the significance of not disclosing banking details to the public:

> *You're not going to give out any information regarding bank accounts. If people want to take advantage of you, they're gonna do it anyways.*

Stigmatization

This theme discusses concerns about how the use of Community ASAP could lead to stigma for people living with dementia. People living with dementia may experience stigma due to their dementia diagnosis and assumptions about their abilities and behaviors. People may assume that individuals living with dementia do not have the capacity to make decisions for themselves. Therefore, sharing personal information such as a cognitive impairment or dementia diagnosis publicly could heighten the existing risk of stigmatization for people living with dementia. A police officer (first responder 3) with experience with missing persons expressed the following:

The one thing we are always cognizant of in our media releases, is our management and our upper management are always cognizant of the...stigma.

A service provider also explained the stigma associated with using alert systems:

> *So, it's such a stigma, it's such a huge issue that sometimes you might think that having an alert system would be really useful for the community to know how to interact or to keep their eye out for somebody with dementia in the community. But is a stigma sort of perpetuating.* [Service provider 7, a service provider who works directly with people living with dementia providing education and counseling]

Participants pointed out that the public may not understand how to approach people living with dementia or communicate with them, hence the need to address dementia-related stigma and misconceptions through public education and awareness campaigns. It was also noted that service providers and first responders should receive adequate training on recognizing, communicating with, and responding to missing persons with dementia. A service provider discussed the importance of using nonstigmatizing language when interacting with missing persons with dementia:

> *I think one of the issues we may see is how responsive behaviors are presented. So, if there are triggers, because I know it is important to communicate that to the community, like you know do not touch this person because they might respond in this way. How can we train coordinators to use language that is going to be non-stigmatizing but still express the nature of that behavior.* [Service provider 1, a service provider who worked directly with people living with dementia providing education and counseling]

Informed and Knowledgeable Consent

This theme focuses on concerns related to ensuring individuals' informed and knowledgeable consent when using Community ASAP. Participants highlighted the importance of people living with dementia having the capacity to understand the implications of an alert system, its objectives, and the reasons behind collecting and sharing their personal data. They emphasized the necessity of providing sufficient information about Community ASAP to those living with dementia and respecting their autonomy in decision-making. However, dementia can affect an individual's ability to process and understand information adequately, leading to challenges in decision-making regarding alert systems, which does not always align with a simple *present or absent* perspective. This concern was elaborated on by a service provider:

> *I think the idea of capacity and the ability to give informed consent is not an on/off button. And to always have that conversation as much as possible to the extent that it is possible for you and whether that's through care partners. I think it's really at the heart of it, really enabling that person to have the choice. But the challenge with that is that on one day the person might say "yeah I'm completely fine with it" and then the next day they're like "ugh, I'm not fine with it at all."* [Service provider 7, a service provider who worked directly with people living with dementia providing education and counseling]

For individuals with limited decision-making capacity, service providers should seek consent from their proxies or substitute decision makers (eg, relatives) and make decisions in their best interests. Participants emphasized the importance of a respectful decision-making process that considers the evolving wishes and preferences of individuals with dementia as the condition advances. A first responder discussed the significance of informed consent and ensuring that decisions made on behalf of an individual living with dementia align with their preferences:

> *Informed consent is the only consent and that would certainly help alleviate some problems down the road, and I would think if the person can't give consent, because of their mental capacity, hopefully the caregiver would have that Power of Attorney to be able to make that decision for them, on their behalf. What's the end goal? The end goal is to try and help that person the best you can, and if they are in a position that is maybe not good for them, at least hopefully we put enough checkmarks in place that they'll come out okay, and that's what you have to look at.* [First responder 2, an experienced police officer and search and rescue expert]

The need to involve people living with dementia in the decision-making process regarding the use of Community ASAP was discussed by participants. It was emphasized that involving the person with dementia in decisions at an early stage is crucial as dementia progression can greatly impact decision-making capacity. Care partner 5, a family member of a person living with dementia who had experience with locator devices, expressed the following:

> *You need to be respectful and try to include their voice as much as possible, knowing that cognitively that could change at any time, and if it does...I would hope that the person has the best interests for that individual.*

Discussion

Principal Findings

Our study aimed to identify the implications of public disclosure of personal information in an alert system called Community ASAP for people living with dementia who are at risk of going missing. The implications were discussed under 4 key themes: *right to autonomy*, *safety versus privacy*, *stigmatization*, and *informed and knowledgeable consent.*

Community ASAP differs from public alert systems for missing persons with dementia such as Silver Alerts. While Community ASAP is a subscription-based, localized alert system that provides real-time updates and geofencing features, Silver Alerts can be government funded or community funded and use large-scale channels such as highway signs and media to reach a wider audience [3,6]. Given the limited research on subscription-based, localized alert systems for missing persons with dementia, our references drew primarily from studies on Silver Alerts.

Participants raised concerns about respecting individuals' right to autonomy when using Community ASAP. The literature emphasizes the importance of honoring the autonomy of people

living with dementia in alert systems [17,32]. Autonomy, or self-determination, is a fundamental human right tied to a person's capacity to assess options, weigh risks, communicate choices, and act accordingly [33,34]. Unlike children requiring guardianship, adults with dementia, even those with appointed guardians, retain some decision-making capacity [33].

People living with dementia may go missing intentionally to avoid disclosing their location or unintentionally due to disorientation or difficulty recognizing familiar places [16,35]. Respecting the right to go missing intentionally is controversial as risk is often viewed negatively and people living with dementia are considered a vulnerable group needing protection [36,37]. The use of technology to locate missing persons raises ethical concerns about balancing autonomy and privacy with safety [6,17,34]. For example, dementia-related wandering is typically seen as "risky" due to potential harm, but it can also represent autonomy and agency and even provide health benefits as a form of exercise [38]. In care settings, health care providers may monitor "wandering" due to ethical responsibilities and potential litigation fears [39]. However, disclosing personal information can violate a person's autonomy [6]. It is crucial to balance safety with respect for autonomy, necessitating open discussions between people living with dementia and their care partners or health care professionals [39]. These conversations should focus on the person's values, care preferences, and ways to support their choices while living with risks.

The public disclosure of personal information in Community ASAP also raises concerns about balancing safety and privacy. While safety focuses on locating missing persons and minimizing risks of harm, privacy emphasizes the right to choose what information is shared with others [33]. Participants expressed that disclosing details such as dementia diagnoses could violate privacy rights. Similarly, Wasser and Fox [32] found in their analysis of Silver Alerts that such systems can unintentionally infringe on the privacy rights of individuals with dementia. In addition, cultural beliefs influence perceptions of dementia [40] and privacy as some cultures view such diagnoses as private. Understanding these cultural factors is essential when using alert systems while ensuring the privacy of people living with dementia.

For most participants, the safety of missing persons outweighs privacy rights. The literature shows that care partners often prioritize safety over privacy concerns [15,16]. Alert systems can serve as a safety net for people living with dementia, providing comfort and independence while easing care partners' worries. Balancing safety and privacy requires understanding individual preferences and the perceptions of risk. Disclosure of personal information may be necessary for safety, but technology can also create a false sense of security if it fails [8,16,33]. To reduce the risk of individuals going missing, additional measures such as GPS devices and return home interviews should be used alongside alert systems [8]. Informed dialogue among people living with dementia, care partners, and health care professionals is essential, along with effective policies to address these concerns [6,16,33].

Public disclosure of personal information in Community ASAP can contribute to the stigmatization and abuse of people living with dementia. Stigma can lead to delays in seeking help, exclusion from social activities, and reduced quality of life [41,42]. Participants in this study emphasized the need for dementia education and staff training to address stigma and improve communication regarding missing persons. Community education is also crucial to raise awareness about the risks associated with alert systems [6,8]. Furthermore, a participatory approach in the design, development, and evaluation of alert systems is necessary [3].

People living with dementia are at heightened risk of abuse due to their diagnosis [43,44]. Publicly sharing information such as dementia diagnoses and home addresses can lead to financial exploitation and other crimes. For example, the United States' Silver Alert may broadcast home addresses, license plates, and photographs of missing persons, unintentionally increasing risks of victimization or identity theft [7,16,17]. It is vital to limit the information shared in alert systems to what is necessary and implement safety measures such as using general locations instead of specific addresses and ensuring that sensitive data are controlled and removed promptly [3].

Another concern regarding the public disclosure of personal information in Community ASAP is informed and knowledgeable consent. Consent is considered knowledgeable if "the individual knows the purposes of the collection, use or disclosure and knows that they have the right to give, withhold or withdraw consent" [45]. Participants noted that people living with dementia may face challenges in providing informed consent for using alert systems, often due to assumptions about their capacity [6,17]. While dementia can impact decision-making capacity in some cases, individuals can still make decisions with support from family, friends, or health care professionals [33,46]. Assessment of decision-making capacity can be difficult as dementia-related cognitive fluctuations affect attention and memory [46,47]. A dementia diagnosis does not automatically imply incapacity [46], and capacity assessments should account for the person's stage of dementia and be conducted during moments when they are experiencing minimal confusion [47]. Continuous consent should be sought as the condition progresses, and early discussions are encouraged to address future decision-making needs. For those requiring substitute decision makers, involving them in decisions respects autonomy and dignity. Regardless of dementia progression, individuals should be included in conversations about their health and the use of alert systems, ensuring that their voices are considered in the design and implementation of these technologies.

Strengths and Limitations

While there is research on public alert systems such as Silver Alerts, little is known about the implications of public disclosure of personal information in subscription-based, localized alert systems. Our study addressed this gap by examining how Community ASAP may infringe on the rights of people living with dementia. Specifically, it highlights the need to understand and address end users' concerns about disclosure of personal information, the right to autonomy, the balance between safety and privacy, and informed and knowledgeable consent. Including participants with lived and professional experience

with dementia and technologies for managing risks of going missing provided diverse perspectives on a topic with limited existing research. The qualitative descriptive approach used in this study enabled an in-depth exploration of concerns associated with public disclosure of personal information in subscription-based, localized alert systems.

Community ASAP and this study have limitations. The subscription-based, localized design of Community ASAP and the small participant sample size for each category limit its generalizability. The subscription model may exclude lower-income users and create a self-selection bias, whereas the app's reliance on a high density of community users may reduce its effectiveness in underengaged or rural areas. These factors may have influenced participant responses and perspectives. To enhance broader applicability, addressing barriers such as cost, accessibility, and adoption in diverse communities would be essential in future research.

In addition, it was unclear how much experience participants had with localized alert systems. Thus, the concept of the release of personal information in such alert systems was hypothetical. Participants' perspectives may differ with real-world experience, and the abstract nature of the topic may have led to the generation of data that are somewhat limited compared to those that a more concrete topic would yield. Recruitment methods also presented a limitation, potentially privileging certain voices. Participants were predominantly White individuals, which does not represent the demographics of Canada. In addition, participants were recruited through the researchers' networks, likely including individuals who were more vocal, confident, or willing to participate. This may have excluded underrepresented perspectives such as those from ethnic communities where alert systems might be viewed differently or where stigma around dementia is more significant.

Conclusions

Localized alert systems such as Community ASAP differ from existing Silver Alert programs that push alerts to the public through the media. Instead, Community ASAP pushes alerts to targeted community members who volunteer to be extra eyes and ears on the lookout for missing persons with dementia. This can promote community engagement in locating missing persons with dementia. However, the public disclosure of personal information in alert systems has significant implications. Such disclosure can compromise an individual's right to autonomy and privacy and lead to stigmatization. The balance of privacy rights with safety concerns presents a challenge. Informed and knowledgeable consent should be a fundamental part of using alert systems. Regardless of the stage of dementia, individuals living with the condition have a right to be included in conversations about their health, including the use of alert systems. An understanding of these implications paves the way to respecting the rights of people living with dementia at risk of going missing and improving their safety and well-being.

Acknowledgments

This work was supported by the Centre for Aging + Brain Health Innovation and Aging Gracefully Across Environments Using Technology to Support Wellness, Engagement, and Long Life from the Networks of Centres of Excellence of Canada. Community ASAP was conceived by Ron Beleno, a caregiver for his father who lived with dementia. The authors thank the study participants for contributing their time and feedback, including the Calgary Missing Older Adult Resource Network, Alzheimer Society of Ontario, Alzheimer Society of Durham Region, and Coquitlam Search and Rescue, as well as Canadian police services in Durham, Ontario; Toronto, Ontario; and Calgary, Alberta, Canada.

Conflicts of Interest

None declared.

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2024-06-22]
2. Ward R, Rummery K, Odzakovic E, Manji K, Kullberg A, Clark A, et al. Getting lost with dementia: encounters with the time-space of not knowing. Health Place 2022 Nov;78:102940 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102940] [Medline: 36403420]
3. Neubauer N, Daum C, Miguel-Cruz A, Liu L. Mobile alert app to engage community volunteers to help locate missing persons with dementia. PLoS One 2021;16(7):e0254952. [doi: 10.1371/journal.pone.0254952] [Medline: 34280219]
4. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 2016 Apr;12(4):459-509. [doi: 10.1016/j.jalz.2016.03.001] [Medline: 27570871]
5. Buckley DM. Towards more effective missing women investigations: police relationships with victims' families, the community and the media a policy discussion. Missing Women Commission of Inquiry. 2012. URL: https://missingwomen.library.uvic.ca/wp-content/uploads/2010/10/POL-2-March-2012-MB-Towards-More-Effective-Missing-Women-Investigations.pdf [accessed 2023-09-18]
6. Gergerich E, Davis L. Silver alerts: a notification system for communities with missing adults. J Gerontol Soc Work 2017 Apr;60(3):232-244. [doi: 10.1080/01634372.2017.1293757] [Medline: 28409711]
7. Gier VS, Kreiner DS, Lampinen JM. Factors affecting recognition of senior citizens in a silver alert. J Police Crim Psych 2016 Oct 18;32(3):185-196. [doi: 10.1007/S11896-016-9210-0]

8. Adekoya A, Daum C, Neubauer N, McLennan L, Liu L. Policy brief Community alert systems for missing vulnerable adults: a call to action for policy makers. University of Waterloo. URL: https://uwaterloo.ca/aging-innovation-research-program/sites/default/files/uploads/files/policy_brief_community_alert_systems_finalcitation.pdf [accessed 2024-04-29]
9. Emergency alerts and the national public alerting system. Government of Canada. URL: https://crtc.gc.ca/eng/television/services/alert.htm [accessed 2023-09-18]
10. Alert criteria. BC Silver Alert. URL: https://bcsilveralert.ca/ [accessed 2023-09-18]
11. Bill 204: missing persons (silver alert) amendment act. Legislative Assembly of Alberta. URL: https://www.assembly.ab.ca/assembly-business/bills/bill?billinfoid=11993&from=bills [accessed 2023-09-18]
12. the missing persons amendment act (silver alert): 2nd session, 41st legislature. The Legislative Assembly of Manitoba. 2017. URL: https://web2.gov.mb.ca/bills/41-2/b214e.php [accessed 2023-09-18]
13. Bill 74, missing persons amendment act. Legislative Assembly of Ontario. URL: https://www.ola.org/en/legislative-business/bills/parliament-43/session-1/bill-74#:~:text=A%20vulnerable%20persons%20alert%20can,by%20a%20person%20in%20a [accessed 2023-09-18]
14. Quebec to launch 'Silver Alert' pilot project to locate missing seniors. Canadian Broadcasting Corporation (CBC). URL: https://www.cbc.ca/news/canada/montreal/silver-alert-quebec-1.6464902 [accessed 2023-09-18]
15. Carr D, Muschert GW, Kinney J, Robbins E, Petonito G, Manning L, et al. Silver alerts and the problem of missing adults with dementia. Gerontologist 2010 Apr 25;50(2):149-157. [doi: 10.1093/geront/gnp102] [Medline: 19556393]
16. Griffin T, Wiecko F. “Worth it if it saves just one life”? Ethical reservations regarding AMBER alert. Crim Justice Ethics 2015 Aug 14;34(2):156-176. [doi: 10.1080/0731129X.2015.1059572]
17. Petonito G, Muschert GW, Carr DC, Kinney JM, Robbins EJ, Brown JS. Programs to locate missing and critically wandering elders: a critical review and a call for multiphasic evaluation. Gerontologist 2013 Feb;53(1):17-25 [FREE Full text] [doi: 10.1093/geront/gns060] [Medline: 22565495]
18. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
19. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb 30;40(1):23-42 [FREE Full text] [doi: 10.1002/nur.21768] [Medline: 27686751]
20. Neergaard MA, Olesen F, Andersen RS, Sondergaard J. Qualitative description - the poor cousin of health research? BMC Med Res Methodol 2009 Jul 16;9:52 [FREE Full text] [doi: 10.1186/1471-2288-9-52] [Medline: 19607668]
21. Emmel N. Purposeful sampling. In: Emmel N, editor. Sampling and Choosing Cases in Qualitative Research: A Realist Approach. Thousand Oaks, CA: Sage Publications; 2013.
22. Patton MQ. Qualitative Research & Evaluation Methods. Thousand Oaks, CA: Sage Publications; 2002.
23. Adekoya AA, Guse L. Walking interviews and wandering behavior: ethical insights and methodological outcomes while exploring the perspectives of older adults living with dementia. Int J Qual Methods 2020 Apr 20;19(3 Suppl. 2):89-95. [doi: 10.1177/1609406920920135]
24. Brooks J, Savitch N, Gridley K. Removing the 'gag': Involving people with dementia in research as advisers and participants. Soc Res Pract 2017;3:14 [FREE Full text]
25. O'Connor CM, Liddle J, O'Reilly M, Meyer C, Cartwright J, Chisholm M, et al. Advocating the rights of people with dementia to contribute to research: considerations for researchers and ethics committees. Australas J Ageing 2022 Jun 25;41(2):309-313. [doi: 10.1111/ajag.13023] [Medline: 34821448]
26. Holtz LR. Applying ethical principles to enrolling older adults with cognitive impairment in research. In: Byrdsong Q, editor. Sage Research Methods Cases: Medicine and Health. Thousand Oaks, CA: Sage Publications; 2023.
27. Richards L, Morse JM. Readme First for a User's Guide to Qualitative Methods. Thousand Oaks, CA: Sage Publications; 2007.
28. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
29. Barber JP, Walczak KK. Conscience and critic: peer debriefing strategies in grounded theory research. In: Proceedings of the 2009 Conference on Annual Meeting of the American Educational Research Association. 2009 Presented at: AERA '09; April 13-17, 2009; San Diego, CA p. 1-19 URL: https://www.researchgate.net/publication/242479874_Conscience_and_Critic_Peer_Debriefing_Strategies_in_Grounded_Theory_Research [doi: 10.4135/9781412986281.n252]
30. Mayan MJ. Essentials of Qualitative Inquiry. Walnut Creek, CA: Routledge; 2009.
31. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. Int J Qual Methods 2002 Jun 01;1(2):13-22. [doi: 10.1177/160940690200100202]
32. Wasser TD, Fox PK. For whom the bell tolls: Silver Alerts raise concerns regarding individual rights and governmental interests. J Am Acad Psychiatry Law 2013;41(3):421-429. [Medline: 24051596]
33. Liu L, Daum C, Neubauer N, Miguel Cruz A, Rios Rincón A. Autonomy and Independence: Aging in an Era of Technology. New York, NY: Morgan & Claypool Publishers; 2022.

34. Bantry-White E. Supporting ethical use of electronic monitoring for people living with dementia: social work's role in assessment, decision-making, and review. J Gerontol Soc Work 2018 Apr 16;61(3):261-279. [doi: 10.1080/01634372.2018.1433738] [Medline: 29381128]
35. Greene KS, Clarke CL, Pakes F, Holmes L. People with dementia who go missing: a qualitative study of family caregivers decision to report incidents to the police. Policing 2019;13(2):241-253. [doi: 10.1093/police/paz007]
36. Brannelly T. Citizenship and people living with dementia: a case for the ethics of care. Dementia (London) 2016 May 10;15(3):304-314. [doi: 10.1177/1471301216639463] [Medline: 27170583]
37. Robinson L, Hutchings D, Dickinson HO, Corner L, Beyer F, Finch T, et al. Effectiveness and acceptability of non-pharmacological interventions to reduce wandering in dementia: a systematic review. Int J Geriatr Psychiatry 2007 Jan;22(1):9-22. [doi: 10.1002/gps.1643] [Medline: 17096455]
38. Adekoya AA, Guse L. Wandering behavior from the perspectives of older adults with mild to moderate dementia in long-term care. Res Gerontol Nurs 2019 Sep 01;12(5):239-247 [FREE Full text] [doi: 10.3928/19404921-20190522-01] [Medline: 31158296]
39. Adekoya A. Tensions between risk, autonomy, and health care settings. In: Liu L, Daum C, Neubauer N, Miguel Cruz A, Rios Rincón A, editors. Autonomy and Independence: Aging in an Era of Technology. New York, NY: Morgan & Claypool Publishers; 2022:69-70.
40. Hillman A, Latimer J. Cultural representations of dementia. PLoS Med 2017 Mar 28;14(3):e1002274 [FREE Full text] [doi: 10.1371/journal.pmed.1002274] [Medline: 28350889]
41. Benbow SM, Jolley D. Dementia: stigma and its effects. Neurodegen Dis Manage 2012 Apr 16;2(2):165-172. [doi: 10.2217/nmt.12.7]
42. Rewerska-Juśko M, Rejdak K. Social stigma of people with dementia. J Alzheimers Dis 2020;78(4):1339-1343. [doi: 10.3233/JAD-201004] [Medline: 33185610]
43. Dong X, Chen R, Simon MA. Elder abuse and dementia: a review of the research and health policy. Health Aff (Millwood) 2014 Apr;33(4):642-649 [FREE Full text] [doi: 10.1377/hlthaff.2013.1261] [Medline: 24711326]
44. Storey JE. Risk factors for elder abuse and neglect: a review of the literature. Aggress Violent Behav 2020 Jan;50:101339. [doi: 10.1016/j.avb.2019.101339]
45. Elements of consent. Information and Privacy Commissioner of Ontario. URL: https://www.ipc.on.ca/part-x-cyfsa/consent-and-capacity/elements-of-consent/#_edn2 [accessed 2023-09-18]
46. Poppe C, Elger BS, Wangmo T, Trachsel M. Evaluation of decision-making capacity in patients with dementia: challenges and recommendations from a secondary analysis of qualitative interviews. BMC Med Ethics 2020 Jul 06;21(1):55 [FREE Full text] [doi: 10.1186/s12910-020-00498-y] [Medline: 32631304]
47. Trachsel M, Hermann H, Biller-Andorno N. Cognitive fluctuations as a challenge for the assessment of decision-making capacity in patients with dementia. Am J Alzheimers Dis Other Demen 2015 Jun 18;30(4):360-363 [FREE Full text] [doi: 10.1177/1533317514539377] [Medline: 24946784]

Edited by X Xiang; submitted 29.07.24; peer-reviewed by L Haslam-Larmer; comments to author 03.12.24; revised version received 16.12.24; accepted 20.12.24; published 07.02.25.

Please cite as:
Adekoya A, Daum C, Neubauer N, Miguel-Cruz A, Liu L
Implications of Public Disclosure of Personal Information in a Mobile Alert App for People Living With Dementia Who Go Missing: Qualitative Descriptive Study
JMIR Aging 2025;8:e64847
URL: https://aging.jmir.org/2025/1/e64847
doi:10.2196/64847
PMID:39918846

Original Paper

Transcultural Adaptation, Validation, Psychometric Analysis, and Interpretation of the 22-Item Thai Senior Technology Acceptance Model for Mobile Health Apps: Cross-Sectional Study

Nida Buawangpong[1], MD; Penprapa Siviroj[2,3], DrPH; Kanokporn Pinyopornpanish[1], MD, PhD; Wachiranun Sirikul[2,4,5], MSc, MD

[1]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Center of Data Analytics and Knowledge Synthesis for Health Care, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

Corresponding Author:
Wachiranun Sirikul, MSc, MD
Department of Community Medicine
Faculty of Medicine
Chiang Mai University
110 Intrawarorot road
Sriphum, Mueng
Chiang Mai, 50200
Thailand
Phone: 66 53935472
Email: wachiranun.sir@cmu.ac.th

Abstract

Background: The rapid advancement of technology has made mobile health (mHealth) a promising tool to mitigate health problems, particularly among older adults. Despite the numerous benefits of mHealth, assessing individual acceptance is required to address the specific needs of older people and promote their intention to use mHealth.

Objective: This study aims to adapt and validate the senior technology acceptance model (STAM) questionnaire for assessing mHealth acceptance in the Thai context.

Methods: In this cross-sectional study, we adapted the original, 38-item, English version of the STAM using a 10-point Likert scale for mHealth acceptability among the Thai population. We translated the mHealth STAM into Thai using forward and backward translation. A total of 15 older adults and experts completed the pilot questionnaire and were interviewed to assess its validity. The pilot items of the Thai mHealth STAM were then reworded and revised for better comprehension and cross-cultural compatibility. The construct validity of the Thai mHealth STAM was evaluated by a multidimensional approach, including exploratory and confirmatory factor analysis and nonparametric item response theory analysis. Discriminative indices consisting of sensitivity, specificity, and area under the receiver operating characteristic (AUROC) were used to determine appropriate banding and discriminant validity for the intention to use mHealth. Internal consistency was assessed using Cronbach α and McDonald ω coefficients.

Results: Out of the 1100 participants with a mean age of 62.3 (SD 8.8) years, 360 (32.7%) were adults aged 45-59 years, and 740 (67.3%) were older adults aged 60 years and older. Of the 40-item pilot questionnaire, exploratory factor analysis identified 22 items with factor loadings >0.4 across 7 principal components, explaining 91.45% of the variance. Confirmatory factor analysis confirmed that 9-dimensional sets of 22 items had satisfactory fit indices (comparative fit index=0.976, Tucker-Lewis index=0.968, root mean square error of approximation=0.043, standardized root mean squared residual=0.044, and R^2 for each item>0.30). The score banding D (low≤151, moderate 152-180, and high≥181) was preferred as the optimal 22-item Thai mHealth STAM cutoff score based on the highest sensitivity of 89% (95% CI 86.1%-91.5%) and AUROC of 72.4% (95% CI 70%-74.8%) for predicting the intention to use mHealth. The final Thai mHealth STAM, consisting of 22 items, exhibited remarkable internal consistency, as evidenced by a Cronbach α of 0.88 (95% CI 0.87-0.89) and a McDonald ω of 0.85 (95% CI 0.83-0.87). For all 22 items, the corrected item-total correlations ranged between 0.26 and 0.71.

Conclusions: The 22-item Thai mHealth STAM demonstrated satisfactory psychometric properties in both validity and reliability. The questionnaire has the potential to serve as a practical questionnaire in assessing the acceptance and intention to use mHealth among pre-older and older adults.

(JMIR Aging 2025;8:e60156) doi:10.2196/60156

KEYWORDS

STAM; senior technology acceptance model; validity; reliability; mHealth; older adult; technology acceptance; mobile health; app; transcultural adaptation; psychometric analysis; geriatrics; cross-sectional study; Thai; theory analysis; Cronbach α; McDonald ω; quality of life; well-being; social media; telehealth; health informatics; eHealth; mobile phone

Introduction

As the global population ages, the integration of technology into the lives of older adults becomes increasingly crucial for enhancing their quality of life, independence, and well-being [1]. An emerging technology that promotes healthy aging is mobile health (mHealth). mHealth refers to medical and public health services facilitated by mobile devices [2]. It can provide individualized care plans for older adults to sustain functional ability and enhance quality of life [3]. Examples of mHealth innovations for older adults include supporting services for age-friendly health and facilitating the establishment of behavioral changes [3,4]. However, the adoption of technology, for example, mHealth, among older adults remains a complex and multifaceted issue, influenced by various factors such as individual perception and experience, ease of use, technological support, and sociocultural contexts [5,6]. To address this challenge, numerous theoretical frameworks have been proposed to understand and predict older adults' acceptance of technology.

Assessing technology acceptance is essential for the successful implementation and use of mHealth technologies, as it directly influences user engagement, health outcomes, and health care delivery efficiency. Understanding acceptance helps developers create user-friendly applications [7,8], improves health outcomes through better adherence to interventions [9,10], and guides implementation strategies to address barriers effectively [11,12]. It also informs policy makers and administrators, enabling evidence-based decisions on mHealth investments [13,14]. Therefore, the lack of validated questionnaires for assessing technology acceptance could lead to a limited understanding of user needs and missed opportunities for improvement. Addressing this gap by developing and validating robust assessment is critical for maximizing the benefits of mHealth technologies and ensuring their effective adoption across diverse populations.

In the field of mHealth, various instruments and frameworks have been developed to assess adoption, intention to use, and acceptance. Established instruments like the Health Information Technology Usability Evaluation Scale (Health-ITUES) [15,16], System Usability Scale (SUS) [17,18], and mHealth App Usability Questionnaire (MAUQ) [19,20] provided insights focusing on user experiences and satisfaction. Broader frameworks include the unified theory of acceptance and use of technology (UTAUT) [21,22], which was extended to include additional factors relevant to mHealth, such as trust and perceived reliability, and was used in various studies to predict mHealth acceptance; the Fit between Individuals, Tasks, and Technology (FITT) Framework [23] is another, which was introduced to measure acceptance in clinical environments, emphasizing the alignment between user needs and technology capabilities. Despite their usability, these instruments and frameworks often lack specificity when addressing the unique needs of older adults.

The senior technology acceptance model (STAM) [24] stands out due to its tailored approach for older adults, which addresses their unique challenges and enhances the relevance of mHealth technologies for this population, making it more relevant than general models like the technology acceptance model (TAM) [25] or the UTAUT [26]. Furthermore, it emphasizes the role of social influence and support, which are critical for older adults who may rely on family and caregivers for technological adoption and addressing common health conditions in older adults, such as cognitive load and physical limitations.

The STAM was first proposed by Chen and Chan [24] in 2014 and has gained prominence for its focus on the unique needs and characteristics of older adults. This model was developed based on a study of 1012 older adults aged 55 years and older in Hong Kong, and it specifically targets older adults as its primary population of interest. The STAM integrates concepts from established technology acceptance frameworks, such as the TAM [25] and the UTAUT [26], tailored to address the specific considerations of older adults and provides a thorough framework for studying the factors that influence technological adoption in this age group. The study indicated 8 dimensions associated with technology acceptance in older adults, which included gerontechnology self-efficacy, gerontechnology anxiety, facilitating conditions, self-reported health conditions, cognitive ability, social relationships, attitude toward life and satisfaction, and physical functioning. Sociodemographic factors such as age, gender, education, and economic status are taken into account [24].

While the STAM has been used in different cultural contexts in other Asian countries, including Hong Kong [24] and South Korea [27], its applicability to the Thai population has not been validated. Thailand, like many other countries, is experiencing rapid population aging, emphasizing the urgency of understanding and promoting health technology acceptance among older adults [28]. However, cultural background, social norms, and technological infrastructures specific to Thailand may influence older adults' perceptions and behaviors toward technology differently than in other contexts. Therefore, this study aimed to adapt, validate, and define the interpretation of the STAM questionnaire for evaluating the acceptance and intent to use mHealth in Thailand.

Methods

Study Design and Study Population

The cross-sectional study was conducted from August 2022 to July 2023 through a nationwide, web-based survey and a community survey. Eligible criteria for the study were Thai citizens aged 45 years and older on the date of the survey who could read and communicate in the Thai language and had no underlying conditions or diseases that limited their ability to complete the survey or use mHealth apps (eg, dementia, active psychological problems, or severe visual problems). The web-based survey was disseminated through an assortment of social media platforms, including the department websites, Facebook, Line, Twitter (rebranded as X in 2023), and Instagram. The information on community survey setting and recruitment is described in the section below. For the survey data collection, the respondents to both the web-based and community surveys used the Research Electronic Data Capture (REDCap; Vanderbilt University) survey platform to self-complete the questionnaires. REDCap [21,22] is a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture, (2) audit trails for tracking data manipulation and export procedures, (3) automated export procedures for seamless data downloads to common statistical packages, and (4) procedures for data integration and interoperability with external sources. All study data were collected and managed using REDCap tools hosted at the Faculty of Medicine, Chiang Mai University. All respondents provided their informed consent, which was included in the screening questionnaire and study information sheet, before participating in this survey. The study excluded incomplete respondents or participants who spent less than 2 minutes or more than 60 minutes on the survey. This study was reported in accordance with COSMIN (Consensus-Based Standards for the Selection of Health Status Measurement Instruments)

reporting guidelines for studies on measurement properties of patient-reported outcome measures [29].

Community Survey Setting and Recruitment

The community survey was distributed by the investigator team, consisting of medical students and health care personnel at primary care units from 10 subdistricts in Chiang Mai province. To identify eligible participants in the target area, officers from the subdistrict primary care units reviewed periodic health survey data for community-dwelling adults aged 45 years and older. Subsequently, patients' information was verified with the health-promoting hospital databases to exclude individuals with health conditions that impeded survey participation or mHealth use as described above. The subdistrict primary health care teams invited eligible individuals to participate in the study through individual contact by community health care volunteers, community radio announcements, and posters displayed at primary care units.

Ethical Considerations

The ethical consideration of the human subject research was approved by the Institutional Review Board of the Faculty of Medicine, Chiang Mai University (COM-2565-09079). All respondents provided their informed consent, as outlined in the screening questionnaire and study information page, before participating in this survey. For the web-based survey, respondents remained anonymous, and no identification data were recorded. In the case of the community survey, identification data of eligible participants were used solely for recruitment purposes within each target area and were not recorded in either the survey form or the study database. Participants received 100 Thai Baht (US $3) as compensation for answering the questionnaires.

Translation and Adaptation of the Thai mHealth STAM

The original, English, 38-item STAM is a 10-point Likert scale consisting of 10 subscales and 38 items that capture the acceptance of general technology use for the older adult population. The total ranges from 38 to 380 points, with a higher score indicating greater acceptance of technology. The validity and reliability of 38-item STAM have been established on a satisfactory scale in 1012 older adults aged 55 years and older in Hong Kong [24]. The construct validity of the STAM was also evaluated with the confirmatory factor analysis (CFA) and revealed a satisfactory model fit with the proposed structure (comparative fit index [CFI]=0.938, root mean square error of approximation [RMSEA]=0.054, and standardized root mean square residual [SRMR]=0.075). The reliability of each subscale with Cronbach α coefficients ranged from 0.67 to 0.95.

Translation and adaptation of the Thai mHealth STAM was performed in accordance with the second edition of the International Test Commission (ITC) Guidelines for Translating and Adapting Tests [30]. In accordance with the ITC precondition guidelines, permission from the holder of the intellectual property rights relating to the 38-item STAM was obtained before performing any translation and adaptation of the STAM. The forward and backward translation with an expert reconciliation design was performed as recommended by the ITC test development guidelines. Before beginning the forward translation process, we decided to include a new subscale, perceived barriers, in the Thai STAM version due to the findings from the previous scoping review [31] on adopting mobile apps for health-related interventions among older adults. It revealed that barriers to adopting mHealth apps among older adults were the most common topics identified in the included studies. Insufficient technological skills, perceived lack of capability and time, concerns regarding personal data privacy, and trust in mHealth providers were the four items comprising the perceived barriers subscale. Following the translation protocol, the original, English, 38-item STAM was adapted to specify mHealth apps in all items and then forward translated into Thai by a professional translator to ensure accuracy for the target audience. The expert panel, which included a digital health expert (family physician and epidemiologist), 2 gerontology physicians, and a public health expert in community medicine, reviewed the forward translation of the Thai STAM questionnaire to ensure readability and transcultural adaptation. The backward translation was done by another professional translator into English. Then, the expert panel reconciled the backward translation version with the original STAM version.

The investigator's team resolved any discrepancies by reaching a final consensus through discussions with the expert panel. To ensure the face and content validity of the proposed questionnaire, a literature review, an expert review, and public interviews were incorporated into the adaptation of the Thai mHealth STAM. In total, 15 older adults participated in this phase to complete the pilot 40-item Thai STAM. Participants were subsequently interviewed to assess the following: overall questionnaire readability, clarity of instructions and items/response options, comprehension of the questionnaire, and other feedback on each item. Then, the pilot 40-item Thai STAM was reworded and revised as recommended on input from both participants and expert interviews. Finally, the pilot 40-item Thai mHealth STAM was given to a group of 40 older adults to verify its reliability and scale usability.

Sample Size Estimation

The sample size was estimated based on three parameters, which are as follows: (1) a stable structure for an exploratory factor analysis (EFA) based on the rule of thumb, which is 10 cases per question; (2) expected CFI for a CFA based on the structural equation modeling; and (3) expected Cronbach α for the internal consistency of the questionnaire. For the first parameter, according to the rule of thumb, at least 440 respondents, accounting for 10% of the dropout rate, were required for an EFA. To achieve the expected CFI of 0.95 for a CFA, at least 459 respondents, accounting for 10% of the dropout rate, were required based on an average factor loading of 0.60 and an average factor correlation of .30 to ensure a .05 α (type I) error and power of 90% [32]. For testing overall reliability, at least 146 total respondents were required based on expected Cronbach α=0.80 (SD .05), a confidence level of 95%, and a dropout rate of 10% [33]. All sample size estimation was performed by the web-based sample size calculator [34]. Finally, the minimal required sample size for this study was 920, which was divided into 460 each for the EFA and CFA, respectively.

Statistical Analysis

Descriptive Analysis

All statistical analyses were conducted using Stata (version 17.0; StataCorp). A *P* value below .05 indicated statistical significance. Categorical data were presented as frequency and percentage, while continuous data were described using mean (SD). Univariable analysis for comparison was performed as appropriate. The Thai mHealth STAM item scores were summarized with central estimations, measures of variability, floor and ceiling effect, skewness, and kurtosis tests. The overall psychometric properties of the Thai mHealth STAM were evaluated for validity and reliability as follows:

Dimensionality

To explore and reduce the dimensionality of the proposed questionnaire, an EFA was performed using a principal component analysis (PCA). The selection of PCA over common factor analysis was based on its ability to enhance parsimony and aid in the selection of factors for CFA [35]. Communalities were initially evaluated, and then orthogonal rotation with the varimax criteria and oblique rotation with promax criteria of the component was conducted. The Kaiser-Meyer-Olkin (KMO) measure and the Bartlett test of sphericity were conducted to verify the appropriateness of using factor analysis. A KMO value greater than 0.8 [36] and a Bartlett test with a *P* value less than .05 [37] are suggested for assessing sample adequacy and the suitability of the data for factor analysis, respectively. Eigenvalues greater than 1, the cumulative percentage of variance, and the scree plot with the number of factors that explained more than 5% of the variance were used to determine the number of factors to be retained [38,39]. A parallel study was conducted to validate the optimal threshold for the number of included factors [40]. Then, we used the following criteria to evaluate the adequacy of the EFA results. First, each should be saliently loaded with at least three items to ensure reliability and stability. In case a factor contains only 2 items, the expert panel consensus will be reached to ensure that the factor is meaningful based on the context and theoretical basis. Second, each item should load saliently on only 1 factor without complex or cross-loadings. Third, each factor should demonstrate internal consistency reliability ≥0.70. Fourth, all factors should be theoretically meaningful [35,41,42].

Construct Validity

For a CFA, structural equation modeling using a maximum likelihood estimation was performed to assure the factor structure based on the exploratory factor, as described previously. To determine the appropriateness of the proposed model, the specific fit indices were evaluated as follows: RMSEA<0.100, SRMR<0.100, CFI>0.900, and Tucker-Lewis Index (TLI)>0.900 [43-45]. To establish acceptance of the final structure of the final model, the coefficient of determination (R^2) and item-scale correlation (standardized factor loading) should be at least 0.30 and 0.40, respectively. Finally, a nonparametric item response theory (IRT) analysis was done to confirm that the final Thai mHealth STAM had the unidimensional set for the relationship between the latent trait and the responses to the items [46]. The IRT analysis was assessed based on fundamental assumptions, including unidimensionality, local independence, and monotonicity. Loevinger H coefficients (H^s) less than 0.3, between 0.3 and 0.4, and greater than 0.4, as determined by the item traces, correspond to poor, medium, and strong scalability properties, respectively. The monotonicity assumption criterion was determined by a critical value of less than 80.

Discriminant Validity

To determine the discriminant validity of the final questionnaire, the intention to use mHealth, as indicated in the external question, "If there are available mHealth applications for you, do you want to use them? (yes/no)," was used as the anchor-based question. The discriminative indices, including sensitivity, specificity, and area under the receiver operating characteristic (AUROC), were used with the intention of determining the appropriate cutoff scores. The 6 proposed bandings for the Thai mHealth STAM scores are categorized into low, moderate, and high acceptance based on score tertiles. Associations between these bandings and the intention to use mHealth are presented by adjusted odds ratios (aORs) with 95% CI from a multivariable logistic regression adjusted for potential

confounders such as age, gender, education, income, and living alone.

Reliability

To estimate the correlation statistics for reliability, 95% CI using 1000 bootstrap resampling was presented alongside the reported correlation statistics. An internal consistency consisting of Cronbach α and McDonald ω coefficients was calculated for each item of the final questionnaire, as well as the entirety of the final questionnaire, to determine internal consistency, reliability, and the degree to which every item on a scale measures the same construct. The values of at least .70 indicated acceptable reliability of the questionnaire [47]. In addition, the item-total correlations and the corrected item-total correlations between .20 and .80 were considerably acceptable. A subgroup analysis of adults aged 45-59 years and adults aged 60 years and older was also performed, recognizing the importance of understanding the unique health needs and challenges faced by both current older populations and those who will age into this group in the future.

Results

Findings From the Translation and Adaptation of the Thai mHealth STAM

After reviewing the forward translation, the panel of experts decided to remove 2 items from the gerontechnology self-efficacy subscale, as they were redundant with the facilitating condition (FC) subscale (FC1 and FC2). Independent back-translation provided an additional check of the semantic equivalence of the translation. A total of 4 items, including PU2, PEOU2, P4, and P8, were modified based on the backward translation. For face and content validity, we conducted interviews with 15 older adults similar to the target population. Based on participants' feedback, 4 items (FC1, FC2, C4, and P2) were slightly modified for clarity. In addition, 2 gerontology experts suggested rephrasing 2 items (A1 and A2) regarding attitude to aging and life satisfaction due to the sensitive wording. Finally, the 40-item Thai mHealth STAM in the pilot group of 40 older adults indicated acceptable internal consistency (Cronbach α =0.91). The details of the full 40 items (10 dimensions) of the Thai mHealth STAM are presented in Table S1 in Multimedia Appendix 1.

Participant Characteristics

From the total of 1100 participants, the mean age was 62.3 (SD 8.8) years. The majority of participants were female (776/1100, 70.5%). Among the 1100 participants, 360 (32.7%) were adults aged 45-59 years, and 740 (67.3%) were older adults aged 60 years and older. Statistically significant differences in the characteristics between adults and older adults were observed in marital status (P=.003), education levels (P<.001), income (P<.001), underlying diseases (P<.001), and technology experience (P<.001). The characteristics of the participants of the study population are presented in Table 1. The derived data were randomly divided in a 1:1 ratio into 2 datasets in preparation for the EFA and CFA. The characteristics of the participants involved in the EFA and CFA are described in Table S2 in Multimedia Appendix 1.

Table 1. Participant characteristics of the study population.

Characteristics	Total (N=1100)	Adults (n=360)	Older adults (n=740)	*P* value
Age (year), mean (SD)	62.3 (8.8)	52.4 (5.5)	67 (5.5)	<.001
Male, n (%)	324 (29.5)	99 (27.5)	225 (30.4)	.32
Marital status, n (%)				.003
Single	96 (8.7)	43 (11.9)	53 (7.2)	
Married	747 (67.9)	250 (69.4)	497 (67.2)	
Separated, divorced, or widowed	257 (23.4)	67 (18.6)	190 (25.7)	
Education levels, n (%)				<.001
No education	18 (1.6)	1 (0.3)	17 (2.3)	
Primary school	725 (65.9)	160 (44.4)	565 (76.4)	
Secondary school	97 (8.8)	50 (13.9)	47 (6.4)	
High school and vocational training	162 (14.7)	99 (27.5)	63 (8.5)	
Pre-university	11 (1)	6 (1.7)	5 (0.7)	
Bachelor's degree	79 (7.2)	41 (11.4)	38 (5.1)	
Master's degree	8 (0.7)	3 (0.8)	5 (0.7)	
Income (THB[a]), n (%)				
<10,000	948 (86.2)	275 (76.4)	673 (90.9)	<.001
10,001-30,000	138 (12.5)	79 (21.9)	59 (8)	
>30,001	14 (1.3)	6 (1.7)	8 (1.1)	
Living status, n (%)				.91
Alone	108 (9.8)	34 (9.4)	74 (10)	
With family	988 (89.8)	325 (90.3)	663 (89.6)	
With others	4 (0.4)	1 (0.3)	3 (0.4)	
Living area, n (%)				.49
Urban	220 (20)	78 (21.7)	142 (19.2)	
Sub-urban	377 (34.3)	116 (32.2)	261 (35.3)	
Rural	503 (45.7)	166 (46.1)	337 (45.5)	
Had any underlying disease, n (%)	726 (66)	189 (52.5)	537 (72.6)	<.001
Hypertension, n (%)	495 (45)	115 (31.9)	380 (51.4)	<.001
Dyslipidemia, n (%)	375 (34.1)	87 (24.2)	288 (38.9)	<.001
Diabetes mellitus, n (%)	184 (16.7)	55 (15.3)	129 (17.4)	.37
Chronic kidney disease, n (%)	17 (1.5)	3 (0.8)	14 (1.9)	.18
Vision problems, n (%)	612 (55.6)	208 (57.8)	404 (54.6)	.32
Wore glasses or contact lens, n (%)	399 (65.2)	144 (69.2)	255 (63.1)	.13
Hearing problems, n (%)	120 (10.9)	15 (4.2)	105 (14.2)	<.001
Used hearing aids, n (%)	4 (3.3)	0 (0)	4 (3.8)	.44
Had experience using a smartphone or tablet, n (%)	873 (79.4)	332 (92.2)	541 (73.1)	<.001
Had own smartphone, n (%)	843 (76.6)	317 (88.1)	526 (71.1)	<.001
Had own tablet, n (%)	20 (1.8)	12 (3.3)	8 (1.1)	.009
Had experience using the internet, n (%)	784 (71.3)	323 (89.7)	461 (62.3)	<.001
Had experience in using mHealth[b] apps, n (%)	439 (50.3)	205 (61.7)	234 (43.3)	<.001
Intention to use mHealth apps, n (%)	537 (48.8)	220 (61.1)	317 (42.8)	<.001

[a]THB 1=US $0.0296195.

[b]mHealth: mobile health.

Dimensionality

According to the item analysis, we excluded 6 items from the physical function subscale (P2, P3, P4, P5, P7, and P8) due to a floor effect or ceiling effect of >80% (Table S3 in Multimedia Appendix 1). An EFA was conducted using PCA with 34 remaining items. The Bartlett test of sphericity obtained $P<.001$, indicating that the correlation matrix was not random [37]. The KMO statistic was 0.875, well above the minimum standard for conducting factor analysis [36]. Therefore, we determined that the input data were appropriate for EFA. Subsequently, the rotation of principal components was performed using both orthogonal rotation (varimax) and oblique rotation (promax) in an attempt to achieve a simple structure. Given the fact that an oblique rotation is generally recommended by measurement specialists to facilitate the emergence of factor intercorrelations [48-50], almost all social sciences measurements exhibit some degree of correlation [51]. In addition, the correlation matrix for the factors with oblique (promax) rotation indicated that the highest correlation was 0.445 (Table S3 in Multimedia Appendix 1); we thereby determined that the factors were correlated, and hence, oblique rotation was an appropriate approach. The results of parallel analysis (Table S4 in Multimedia Appendix 1) and PCA with or without oblique (promax) rotation all recommended the retention of 7 factors. According to the previous criteria, 2-item factors were identified, including factor 4 (PBR1 and PBR2) and factor 6 (S1 and S2). The internal consistency of seven factors demonstrated Cronbach α of 0.884 with 95% CI (0.875-0.894), which met acceptable thresholds. Within the context and theoretical framework of the STAM [24] and the UTAUT [11,21,26,52], social factors significantly influence behavioral intentions to use technology, particularly in the use of mHealth. Perceived barriers also play a role in determining intentions to use mHealth, as demonstrated by the aforementioned scoping review [31]. The inclusion of factor 4 and factor 6, which represented perceived barriers and social relationships, was considered appropriate. Based on the priori criteria and consensus of the panel experts, the EFA identified 22 candidate items (ATT1, ATT2, PU1, PU2, PU3, PEOU1, PEOU2, PB1, PB2, ANX1, ANX2, FC2, FC4, FC5, H1, H2, H5, C2, C3, C4, S1, and S2) with factor loadings greater than 0.4 that encompassed the 7 factors. The final EFA result is presented in Table 2.

Table 2. Exploratory factor analysis of the final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items	Factor loadings[a]							Communality value
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	
ATT1	0.639[b]	0.096	0.057	0.027	0.040	0.057	0.060	0.824
ATT2	0.631[b]	0.059	0.060	0.036	0.058	0.071	0.057	0.822
PU1	0.899[b]	0.061	0.030	0.065	0.053	0.040	0.032	0.825
PU2	0.921[b]	0.031	0.027	0.081	0.036	0.054	0.099	0.878
PU3	0.919[b]	0.046	0.015	0.079	0.046	0.056	0.072	0.873
PEOU1	0.473[b]	0.188	0.074	0.298	0.055	0.007	0.175	0.677
PEOU2	0.513[b]	0.169	0.108	0.281	0.080	0.019	0.175	0.746
PBR1	0.115	0.203	0.071	0.854[b]	0.025	0.000	0.038	0.795
PBR2	0.141	0.200	0.061	0.855[b]	0.018	0.009	0.069	0.811
PBR3	0.031	0.880[b]	0.025	0.120	–0.017	0.003	0.032	0.806
PBR4	0.073	0.894[b]	0.012	0.109	–0.016	0.023	0.058	0.831
ANX1	0.130	0.720[b]	0.061	0.183	0.015	0.019	0.044	0.734
ANX2	0.107	0.643[b]	0.081	0.231	0.030	0.031	0.055	0.688
FC2	0.324	0.088	0.032	0.103	0.034	0.054	0.579[b]	0.486
FC4	0.380	0.171	0.147	0.237	0.064	–0.009	0.449[b]	0.606
FC5	0.330	0.201	0.020	0.124	0.032	0.079	0.563[b]	0.504
H1	0.074	–0.020	0.172	0.007	0.723[b]	0.100	0.017	0.573
H2	0.092	–0.009	0.129	0.042	0.713[b]	0.097	0.009	0.556
H5	0.127	–0.037	0.297	0.098	0.505[b]	0.112	0.072	0.498
C1	0.009	0.040	0.614[b]	0.081	0.221	0.066	0.026	0.497
C2	0.163	0.086	0.611[b]	0.164	0.177	0.117	0.073	0.582
C3	0.019	0.054	0.693[b]	0.067	0.144	0.129	0.024	0.547
C4	0.043	0.065	0.624[b]	0.068	0.106	0.237	0.024	0.497
S1	0.126	0.040	0.226	0.021	0.147	0.669[b]	0.024	0.548
S2	0.142	0.034	0.162	–0.012	0.116	0.678[b]	0.040	0.532
% of variance[b]	26.41	15.99	12.37	11.77	10.76	7.32	6.83	—[c]
Cumulative % of variance[b]	26.41	42.40	54.77	66.54	77.30	84.60	91.45	—

[a]The extraction method was principal component analysis, with the rotation method by oblique, promax rotation.

[b]Items load on the assigned factor loadings >0.4 are highlighted.

[c]Not applicable.

Construct Validity

From the EFA, the 22 items of the 7-factor Thai mHealth STAM explained 91.45% of the variance. The unidimensionality of each factor (subscale) and the overall models were assessed by analyzing modification indices in the CFA. Of the 7 factors from the EFA, the CFA of each factor (subscale) showed that only 5 factors, consisting of cognitive ability (C2, C3, C4), perceived barriers (PB1 and PB2), facilitating conditions (FC2, FC4, and FC5), self-reported health conditions (H1, H2, and H5), and social relationships (S1 and S2), showed satisfactory information criteria indices of the CFA, as presented Table 3. Factor 1, which included items from the attitude toward using (ATT1 and ATT2), perceived usefulness (PU1, PU2, and PU3),

and perceived ease of use (PEOU1 and PEOU2), did not meet the CFA criteria due to over-factoring issues (Table S5 in Multimedia Appendix 1). Factor 2, combining perceived barriers (PB3 and PB4) and gerontechnology anxiety (ANX1 and ANX2), was unfit according to CFA criteria, with a low CFI (0.799), low TLI (0.698), and high RMSEA (0.306, 90% CI 0.293-0.319). Attempts to combine subscales also did not meet CFA criteria (Table S5 in Multimedia Appendix 1). However, when items were separated as in the original STAM, including attitude toward using (ATT1 and ATT2), perceived usefulness (PU1, PU2, and PU3), perceived ease of use (PEOU1 and PEOU2), and gerontechnology anxiety (ANX1 and ANX2), these separated factors showed a good fit with CFA criteria (Table S5 in Multimedia Appendix 1). Out of the 9 factors from the single latent factor analysis, 5 were 2-item factors. These were kept in the final CFA model because 3 factors (attitude toward using, perceived ease of use, and gerontechnology anxiety) were originally designed as 2-item factors, similar to the original STAM. The perceived barriers and social relationships were also retained because of their contextual relevance, as described above. Finally, the CFA confirmed 9-dimensional sets of 22 items with satisfactory fit indices, as shown in Table 3. The details of the CFAs of evaluated and reevaluated models are described in Table S4 in Multimedia Appendix 1.

A nonparametric IRT analysis also affirmed the unidimensionality, local independence, and monotonicity of the 22-item model with 8 factors (Table S6 in Multimedia Appendix 1). For the scalability, all 22 items of the Thai mHealth STAM had H^s coefficients over 0.4, which indicates medium to strong scalability properties (Table S6 in Multimedia Appendix 1). The correlation among the final 22-item Thai mHealth STAM subscales ranged from 0.040 to 0.685 (Table S7 in Multimedia Appendix 1). The final 22-item Thai mHealth STAM questions, along with the English version and modeling indices, are described in Table 4.

Each item is scored on a 10-point Likert scale from 1 (very unsatisfied or strongly disagree) to 10 (very satisfied or strongly agree), with reverse scaling for perceived barriers and gerontechnology anxiety.

Table 3. Confirmatory factor analysis of the final Thai mobile health (mHealth) senior technology acceptance model (STAM).

Factor	Number of items	Threshold for acceptable fit					Model fit
		CFI[a] (>0.90)	TLI[b] (>0.90)	RMSEA[c] (<0.10 [90% CI])	SRMR[d] (<0.10)	R^2 (>0.30)	
Attitude toward using	2 items (ATT1 and ATT2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Perceived of benefits	3 items (PU1, PU2, and PU3)	1.000	1.000	<0.001 (<0.001 to <0.001	<0.001	All >0.30	Acceptable
Perceived ease of use	2 items (PEOU1 and PEOU2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Perceived of barriers	2 items (PB1 and PB2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Gerontechnology anxiety	2 items (ANX1 and ANX2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Facilitating conditions	3 items (FC2, FC4, and FC5)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Self-reported health conditions	3 items (H1, H2, and H5)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Cognitive ability	3 items (C2, C3, and C4)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Social relationships	2 items (S1 and S2)	1.000	1.000	<0.001 (<0.001 to <0.001)	<0.001	All >0.30	Acceptable
Final Thai mHealth STAM 9-dimensional model	22 items	0.976	0.968	0.043 (0.039 to 0.047)	0.044	All >0.30	Acceptable

[a]CFI: comparative-fit index.

[b]TLI: Tucker-Lewis Index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean squared residual.

Table 4. The final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items and questions		Mean (SD)	Ceiling, %	Floor, %	Skewness	Kurtosis	Standardized factor loading (95% CI)	R^2
Attitude toward use								
ATT1	Using mobile health applications is a good idea.	8.18 (2.30)	48.45	2.50	–1.11	3.45	0.94 (0.92-0.95)	0.875
ATT2	You like the idea of using mobile health applications.	8.02 (2.39)	45.85	3.11	0.34	2.25	0.93 (0.91-0.95)	0.862
Perceived usefulness								
PU1	Using mobile health applications would enhance your effectiveness in life.	7.66 (2.59)	40.91	4.09	–0.92	2.93	0.91 (0.90-0.92)	0.831
PU2	Using mobile health applications would make your life more convenient.	7.72 (2.61)	43.08	4.09	–0.95	2.93	0.94 (0.93-0.95)	0.897
PU3	You would find mobile health applications useful in your life.	7.82 (2.62)	45.42	3.92	–1.03	3.03	0.94 (0.93-0.95)	0.891
Perceived ease of use								
PEOU1	You would find mobile health applications are easy to use.	6.37 (3.18)	28	11.64	–0.34	1.76	0.83 (0.81-0.86)	0.693
PEOU2	You could be skillful at using mobile health applications.	6.93 (3.06)	34	9	–0.62	2.10	0.89 (0.87-0.92)	0.803
Gerontechnology anxiety								
ANX1	You feel apprehensive about using mobile health applications.	5.89 (3.16)	23.56	13.2	–0.09	1.70	0.89 (0.86-0.92)	0.799
ANX2	You hesitate to use the technology for fear of making mistakes you cannot correct.	5.77 (3.11)	20.66	13.92	–0.38	1.75	0.95 (0.93-0.98)	0.918
Perceived barriers								
PB1	You need to put in a lot of effort to use mobile health applications?	4.81 (3.14)	13.27	21.42	0.34	1.79	0.89 (0.86-0.92)	0.704
PB2	You need to spend a lot of time to use mobile health applications?	5.00 (3.19)	14.78	20.35	0.26	1.70	0.96 (0.93-0.98)	0.819
Facilitating conditions								
FC2	A specific person (or group) is available for assistance with difficulties using mobile health applications.	7.45 (3.09)	44.79	10.06	–0.96	2.57	0.63 (0.59-0.68)	0.407
FC4	When you want or need to use mobile health applications, they are accessible to you.	7.20 (2.93)	36.92	7.83	–0.75	2.38	0.79 (0.76-0.83)	0.631
FC5	Your family and friends think/support that you should use mobile health applications.	6.72 (3.33)	36.62	14.04	–0.55	1.83	0.68 (0.64-0.72)	0.461
Self-reported health conditions								
H1	How are your general health conditions?	7.73 (1.75)	21.78	0.27	–0.52	2.79	0.80 (0.77-0.84)	0.651
H2	How are your health conditions compared with the same-age groups?	7.91 (2.03)	33.69	0.18	–0.69	2.54	0.77 (0.73-0.81)	0.595
H5	How well are you able to move around?	8.56 (1.96)	52.85	0.36	–1.37	4.24	0.61 (0.56-0.65)	0.370
Cognitive ability								
C2	How satisfied are you with your ability to learn new information?	7.89 (2.25)	38.38	1.16	–0.98	3.27	0.65 (0.61-0.70)	0.432
C3	How well are you able to concentrate?	8.75 (1.72)	54.23	0.18	–1.51	4.96	0.75 (0.71-0.79)	0.563

Items and questions		Mean (SD)	Ceiling, %	Floor, %	Skewness	Kurtosis	Standardized factor loading (95% CI)	R^2
C4	How satisfied are you with your ability to make decisions?	8.91 (1.58)	58.15	0.09	–1.48	4.64	0.76 (0.72-0.80)	0.578
Social relationships								
S1	How satisfied are you with your personal relationships?	9.34 (1.29)	70.99	0.09	–2.36	8.95	0.84 (0.80-0.89)	0.712
S2	How satisfied are you with the support you get from your friends and family?	9.39 (1.27)	74.01	0.09	–2.59	10.47	0.75 (0.71-0.80)	0.568
Overall (possible range 22-220)		164.16 (30.55)	—[a]	—	—	—	—	0.999

[a]Not applicable.

Discriminant Validity

Considering the absence of a reference standard, it is theoretically reasonable that more participants with higher STAM scores will result in greater acceptance and adoption of technology. The discriminative indices, including sensitivity, specificity, and AUROC, were used to determine the cutoff scores for the proposed questionnaire, considering the intention to use mHealth from the external question. The 6 proposed sets of the final 22-item Thai mHealth STAM bands were classified into low, moderate, and high acceptance, as presented in Table 5. The set D of the possible banding was preferred as the optimal 22-item Thai mHealth STAM cutoff score based on the highest sensitivity of 89% (95% CI 86.1%-91.5%) and AUROC of 72.4% (95% CI 70%-74.8%). This finding also confirmed the discrimination performance of the 22-item Thai mHealth STAM in identifying persons with and without the intention to use mHealth. For set D, low, moderate, and high scores are defined as ≤151, 152-180, and ≥181, respectively. In addition, we conducted a subgroup analysis based on age groups: pre-older adults (aged 45-59 years) and older adults (aged 60 years and older). The result revealed that the set D banding had robust discriminant validity in older adults (AUROC 73%, 95% CI 70%-76%), but the discriminant validity decreased in the pre-older adult group (AUROC 67.7%, 95% CI 63.3%-71.9%). The discriminant validity of the 22-item Thai mHealth STAM by the subpopulation cohorts is shown in Table S8 in Multimedia Appendix 1.

Table 5. Proposed sets of the final 22-item Thai mobile health (mHealth) senior technology acceptance model (STAM) bands.

Possible bandings[a]		Discriminant validity (intention to use mHealth)			
Set and band	Score	aOR[b] (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	AUROC[c] (95% CI)
Set A					
Low	≤121	Ref[d]	Ref	Ref	Ref
Moderate	122-150	2.69[e] (1.21-5.95)	98.5 (97.1-99.4)	16.0 (13.1-19.3)	57.3 (55.6-58.8)
High	≥151	15.53[e] (7.31-5.95)	89.0 (86.1-91.5)	53.5 (49.2-57.6)	71.2 (68.8-73.7)
Set B					
Low	≤131	Ref	Ref	Ref	Ref
Moderate	132- 160	4.37[e] (2.42-7.88)	97.2 (95.4-98.4)	26.1 (22.5-29.9)	61.7 (59.7-63.6)
High	≥161	15.53[e] (8.75-27.52)	78.0 (74.3-81.5)	64.8 (60.7-68.8)	71.4 (68.8-74.1)
Set C					
Low	≤141	Ref	Ref	Ref	Ref
Moderate	142- 170	4.33[e] (2.82-6.66)	93.7 (91.3-95.6)	40.9 (36.8-45.0)	67.3 (65.0-69.5)
High	≥171	13.18[e] (8.-20.26)	66.1 (61.9- 70.1)	75.8 (72.1-79.3)	71.0 (68.3-73.6)
Set D					
Low	≤151	Ref	Ref	Ref	Ref
Moderate	152-180	5.73[e] (4.01–8.19)	89.0 (86.1-91.5)	55.8 (51.6-59.9)	72.4 (70.0-74.8)
High	≥181	12.49[e] (8.45-18.47)	49.9 (45.6-54.2)	84.7 (81.5–87.6)	67.3 (64.7-69.9)
Set E					
Low	≤161	Ref	Ref	Ref	Ref
Moderate	162-190	3.55[e] (2.62-4.83)	76.7 (72.9-80.2)	65.5 (61.5-69.5)	71.1 (68.5-73.4)
High	≥191	8.46[e] (5.74-12.47)	37.4 (33.3-41.7)	90.9 (88.3-93.2)	64.2 (61.8-66.6)
Set F					
Low	≤171	Ref	Ref	Ref	Ref
Moderate	172-200	4.24[e] (3.14-5.74)	65.2 (61.0-69.2)	77.1 (73.4-80.5)	71.1 (68.4-73.8)
High	≥201	7.59[e] (4.68-12.29)	21.8 (18.4-25.5)	95.6 (93.5-97.1)	58.7 (56.7-60.6)

[a]The final 22-item Thai mHealth STAM is highlighted.

[b]aOR: adjusted odds ratio.

[c]AUROC: area under receiver operating characteristic curve.

[d]Ref: reference.

[e]All reported aORs were statistically significant with *P* value<.05. aORs were estimated using a multivariable logistic regression with adjustment for age, gender, education levels (no, primary, secondary, and university education), income levels (low: <10,000 baht, moderate: 10,000-30,000 baht, and high: >30,000 bath), and living alone.

Scale Reliability

Out of 1100 overall participants, the final 22-item Thai mHealth STAM demonstrated an excellent internal consistency in both the Cronbach α (0.88, 95% CI 0.87-0.89) and the McDonald ω coefficients (0.85, 95% CI 0.83-0.87), as shown in Table 6. By subpopulation, the Cronbach α and the McDonald ω coefficients were 0.88 (95% CI 0.86-0.90) and 0.84 (95% CI 0.81- 0.89) for adults aged 45-59 years and 0.88 (95% CI 0.86- 0.89) and 0.83 (95% CI 0.81-0.86) for older adults. All 22 items revealed the corrected item-total correlations ranging from 0.26 to 0.71, achieving a level of acceptance between 0.20 and 0.80 (Table 6).

Table 6. Reliability of the final 22-Item Thai mobile health (mHealth) senior technology acceptance model (STAM).

Items	n	Item-total correlations	Corrected item-total correlations	Average interitem correlation	Cronbach α	McDonald ω
ATT1	1158	0.652	0.605	1.682	0.877	0.879
ATT2	1156	0.652	0.603	1.675	0.877	0.879
PU1	1149	0.684	0.634	1.645	0.875	0.878
PU2	1149	0.720	0.673	1.631	0.874	0.876
PU3	1147	0.713	0.667	1.632	0.874	0.877
PEOU1	1143	0.721	0.664	1.585	0.873	0.878
PEOU2	1144	0.761	0.713	1.575	0.871	0.876
PB1	1130	0.514	0.431	1.682	0.881	0.885
PB2	1130	0.543	0.462	1.666	0.880	0.884
ANX1	1129	0.472	0.385	1.701	0.883	0.886
ANX2	1128	0.481	0.396	1.698	0.883	0.886
FC2	1123	0.546	0.469	1.671	0.880	0.884
FC4	1124	0.703	0.649	1.613	0.874	0.879
FC5	1125	0.586	0.507	1.640	0.879	0.883
H1	1125	0.316	0.263	1.805	0.885	0.888
H2	1125	0.342	0.281	1.790	0.884	0.887
H5	1124	0.401	0.344	1.775	0.883	0.886
C2	1123	0.522	0.465	1.725	0.880	0.883
C3	1123	0.340	0.288	1.800	0.884	0.887
C4	1123	0.366	0.320	1.799	0.883	0.886
S1	1117	0.343	0.305	1.814	0.884	0.886
S2	1116	0.329	0.291	1.817	0.884	0.887
Test scale, Cronbach α (95% CI)	—[a]	—	—	1.701	0.884 (0.875-0.894)	0.85 (0.83-0.87)

[a]Not applicable.

Discussion

Principal Findings

The study aimed to adapt and validate the STAM questionnaire for assessing mHealth technology acceptance among pre-older and older populations regarding the use of health support. The results confirmed the scale's factor structure, supported an 8-factor model with 22 items, and showed good discriminant validity in predicting mHealth intention. The optimal version was a 22-item Thai mHealth STAM using the scoring cutoff (≥152). Subgroup analysis indicated no significant difference in discriminant validity between pre-older and older adults. The scale demonstrated strong internal consistency and stability, with reliability confirmed by Cronbach α and McDonald ω coefficients. This adapted 22-item version is more relevant for assessing mHealth intention among older adults and is suitable for public surveys and routine practice, which take less than 15 minutes to complete.

Our findings are consistent with the previous study conducted by the owner of the original STAM [53], which was subsequently developed into a brief form to save administration time and reduce the burden on respondents. The 14-item brief version of the STAM questionnaire consisted of a 4-factor structure: attitudinal beliefs, control beliefs, gerontechnology anxiety, and health. These findings are consistent with ours, reflecting the original STAM model constructs and the age-related health characteristics of older adults. We observed a decrease in discriminant validity within the pre-older adult group, indicating a need for additional factors to explain their behavioral intentions. For example, older adults with different genders, education levels, income, marital status, and ethnicity may have different intentions and purposes to use mHealth for their health [54].

Strengths and Limitations

On the strength side, this is the first Thai version of the STAM questionnaire suitable for evaluating technology acceptance in Thai older adults. The 22-item Thai STAM version demonstrates structural balance, reliability, and validity in assessing technology acceptance among older individuals. The evaluation process is time-efficient. In addition, this tool can be used with both pre-older adults and older adults to prepare them for engaging with technology in their future lives. Furthermore, the

evaluation takes into account the influence of Thai cultural norms on the adoption and acceptance of mHealth.

However, there are some limitations to consider. Although the psychometric properties of the 22-item Thai mHealth STAM are satisfied through transcultural adaptation in terms of validity and reliability in both the pre-older and older populations, this scale can be applied for use in a broad. However, our study participants may not be representative of the overall Thai pre-older and older populations, as almost all of the participants lived in the northern part of Thailand, particularly in Chiang Mai province. In order to address this concern, future studies, including those based on different regions of Thailand and other specific populations (eg, teenagers, vulnerable groups, minorities, and specific groups of patients) that could potentially derive advantages from mHealth usage, are recommended to expand the generalizability and usability of this scale. Finally, the 22-item Thai mHealth STAM was evaluated based on the board's definition of mHealth. It is possible that the proposed questionnaire may not be compatible with all of the existing mHealth technologies due to the diverse range of mHealth technologies in health care. The patient's choice may vary depending on several factors, such as health care providers, types of services, or the specific application. Hence, we suggest using this questionnaire to assess their acceptance and intention to use it in conjunction with the designated mHealth technology.

Practical Implications of the 22-Item Thai mHealth STAM

The 22-item Thai mHealth STAM offers a practical assessment of patients' acceptability—a crucial factor often overlooked, as evidenced by a recent systematic review of technology acceptability in health care, which revealed that only 10% (142/1219) of the reviewed studies examined patient acceptance [55]. This publicly available questionnaire has the potential to support health care professionals, policy makers, and developers in making informed decisions [56,57], particularly regarding the adoption and acceptance of mHealth within Thai cultural norms. This questionnaire can be incorporated into the research and development (R&D) processes of mHealth and used as a questionnaire to define the target population based on levels of acceptability, as well as ascertain the factors that encourage or hinder the adoption of their mHealth technologies [58,59]. This information is important for informing stakeholders and developers in advance of the mHealth R&D and implementation stages, which necessitate user data for resource allocation and planning in consideration of user requirements and experiences [60,61].

Conclusion

The increasing number of older people, along with their growing adoption of technology, indicates that mHealth technologies might offer a new approach to enhancing the health of older adults with lower health care expenses. Although there are many advantages to using mHealth apps, it is important to consider their acceptance and intention to use them for health-related objectives. We proposed the 22-item Thai mHealth STAM as the questionnaire to evaluate the levels of acceptability and intention to use mHealth in the Thai community of pre-older and older adults. The 22-item Thai mHealth STAM has demonstrated satisfactory psychometric properties in terms of validity and reliability. As a result, it is now feasible to use this questionnaire in a public survey to support stakeholders in making informed decisions. Nevertheless, to improve generalizability and long-term use, further study is needed to investigate the various demographic groups with the specific mHealth interventions.

Acknowledgments

This work was supported by the Faculty of Medicine, Chiang Mai University (grant no. 124-2565), and Chiang Mai University. We would like to express our gratitude to Ms Asapaporn Nimmannit from the Research Institute for Health Sciences at Chiang Mai University for her support in translating the questionnaire. We used QuillBot [62] to review grammar and edit language. We critically reviewed the feedback generated by QuillBot and, based on this, revised the writing using our own words and expressions.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the institutional policies for protecting participant confidentiality but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File , 72 KB - aging_v8i1e60156_app1.docx]

References

1. Lee C. Technology and aging: the jigsaw puzzle of design, development and distribution. Nat Aging 2022;2(12):1077-1079. [doi: 10.1038/s43587-022-00325-6] [Medline: 37118534]

2. Wang Q, Liu J, Zhou L, Tian J, Chen X, Zhang W, et al. Usability evaluation of mHealth apps for elderly individuals: a scoping review. BMC Med Inform Decis Mak 2022;22(1):317 [FREE Full text] [doi: 10.1186/s12911-022-02064-5] [Medline: 36461017]
3. Liaw ST, Ansari S, Jonnagaddala J, Narasimhan P, Ashraf MM, Harris-Roxas B, et al. Use of mHealth for promoting healthy ageing and supporting delivery of age-friendly care services: a systematic review. Int J Integr Care 2019;19(4):147. [doi: 10.5334/ijic.s3147]
4. Changizi M, Kaveh MH. Effectiveness of the mHealth technology in improvement of healthy behaviors in an elderly population-a systematic review. Mhealth 2017;3:51 [FREE Full text] [doi: 10.21037/mhealth.2017.08.06] [Medline: 29430455]
5. Lee C, Coughlin JF. PERSPECTIVE: older adults' adoption of technology: an integrated approach to identifying determinants and barriers. J of Product Innov Manag 2014;32(5):747-759 [FREE Full text] [doi: 10.1111/jpim.12176]
6. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. Computers in Human Behavior 2013;29(4):1715-1724 [FREE Full text]
7. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
8. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data Is not enough: a qualitative exploration of the role of healthcare partnerships and mHealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016;4(1):e5 [FREE Full text] [doi: 10.2196/mhealth.4814] [Medline: 26792225]
9. Bokolo AJ. Application of telemedicine and eHealth technology for clinical services in response to COVID-19 pandemic. Health Technol (Berl) 2021;11(2):359-366 [FREE Full text] [doi: 10.1007/s12553-020-00516-4] [Medline: 33469474]
10. Stover A, Hayes A, Barrera L, Golbeck E, Malloy C. A comparison of use of the mobile electronic health record by medical providers based on clinical setting. J Emerg Invest 2023. [doi: 10.59720/22-196]
11. Garavand A, Samadbeik M, Nadri H, Rahimi B, Asadi H. Effective factors in adoption of mobile health applications between medical sciences students using the UTAUT model. Methods Inf Med 2019;58(4-05):131-139. [doi: 10.1055/s-0040-1701607] [Medline: 32170717]
12. Koivumäki T, Pekkarinen S, Lappi M, Väisänen J, Juntunen J, Pikkarainen M. Consumer adoption of future myData-based preventive eHealth services: an acceptance model and survey study. J Med Internet Res 2017;19(12):e429 [FREE Full text] [doi: 10.2196/jmir.7821] [Medline: 29273574]
13. Agarwal S, Perry HB, Long LA, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
14. Kim S, Lee KH, Hwang H, Yoo SY. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016;16(1):12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
15. Muro-Culebras A, Escriche-Escuder A, Martin-Martin J, Roldán-Jiménez C, De-Torres I, Ruiz-Muñoz M, et al. Tools for evaluating the content, efficacy, and usability of mobile health apps according to the consensus-based standards for the selection of health measurement instruments: systematic review. JMIR Mhealth Uhealth 2021;9(12):e15433 [FREE Full text] [doi: 10.2196/15433] [Medline: 34855618]
16. Schnall R, Cho H, Liu J. Health information technology usability evaluation scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR Mhealth Uhealth 2018;6(1):e4 [FREE Full text] [doi: 10.2196/mhealth.8851] [Medline: 29305343]
17. Hajesmaeel-Gohari S, Khordastan F, Fatehi F, Samzadeh H, Bahaadinbeigy K. The most used questionnaires for evaluating satisfaction, usability, acceptance, and quality outcomes of mobile health. BMC Med Inform Decis Mak 2022;22(1):22 [FREE Full text] [doi: 10.1186/s12911-022-01764-2] [Medline: 35081953]
18. Patel B, Thind A. Usability of mobile health apps for postoperative care: systematic review. JMIR Perioper Med 2020;3(2):e19099 [FREE Full text] [doi: 10.2196/19099] [Medline: 33393925]
19. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
20. Mustafa N, Safii NS, Jaffar A, Sani NS, Mohamad MI, Abd Rahman AHA, et al. Malay version of the mHealth app usability questionnaire (M-MAUQ): translation, adaptation, and validation study. JMIR Mhealth Uhealth 2021;9(2):e24457 [FREE Full text] [doi: 10.2196/24457] [Medline: 33538704]
21. Schomakers EM, Lidynia C, Vervier LS, Valdez AC, Ziefle M. Applying an extended UTAUT2 model to explain user acceptance of lifestyle and therapy mobile health apps: survey study. JMIR Mhealth Uhealth 2022;10(1):e27095. [doi: 10.2196/27095] [Medline: 35040801]
22. Schretzlmaier P, Hecker A, Ammenwerth E. Extension of the unified theory of acceptance and use of technology 2 model for predicting mHealth acceptance using diabetes as an example: a cross-sectional validation study. BMJ Health Care Inform 2022;29(1):e100640. [doi: 10.1136/bmjhci-2022-100640] [Medline: 36379608]

23. Noblin A, Shettian M, Cortelyou-Ward K, Dugre JS. Exploring physical therapists' perceptions of mobile application usage utilizing the FITT framework. Informatics for Health and Social Care 2016;42(2):180-193. [doi: 10.1080/17538157.2016.1178118]
24. Chen K, Chan AHS. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319 [FREE Full text] [doi: 10.2307/249008]
26. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
27. Shin HR, Um SR, Yoon HJ, Choi EY, Shin WC, Lee HY, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. J Med Internet Res 2023;25:e41935 [FREE Full text] [doi: 10.2196/41935] [Medline: 37036760]
28. Jantavongso S. Toward global digital literate citizens: a case of 's aging generation. E J Info Sys Dev Countries 2021;88(2):e12207 [FREE Full text] [doi: 10.1002/isd2.12207]
29. Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res 2021;30(8):2197-2218. [doi: 10.1007/s11136-021-02822-4] [Medline: 33818733]
30. The ITC guidelines for translating and adapting tests (Second edition). International Test Commission. 2017. URL: https://www.intestcom.org/ [accessed 2025-12-10]
31. Ahmad NA, Mat Ludin AF, Shahar S, Mohd Noah SA, Mohd Tohit N. Willingness, perceived barriers and motivators in adopting mobile applications for health-related interventions among older adults: a scoping review. BMJ Open 2022;12(3):e054561 [FREE Full text] [doi: 10.1136/bmjopen-2021-054561] [Medline: 35264349]
32. Kim KH. The relation among fit indexes, power, and sample size in structural equation modeling. Structural Equation Modeling: A Multidisciplinary Journal 2005;12(3):368-390. [doi: 10.1207/s15328007sem1203_2]
33. Bonett DG. Sample size requirements for testing and estimating coefficient alpha. J Educat Behav Stat 2002;27(4):335-340. [doi: 10.3102/10769986027004335]
34. Arifin WN. Sample size calculator (web). 2017-2024 [cited 2022]. 2017. URL: http://wnarifin.github.io [accessed 2025-02-25]
35. Watkins MW. Exploratory factor analysis: a guide to best practice. J Black Psychol 2018;44(3):219-246. [doi: 10.1177/0095798418771807]
36. Kaiser HF. An index of factorial simplicity. Psychometrika 2025;39(1):31-36. [doi: 10.1007/BF02291575]
37. Bartlett MS. A note on the multiplying factors for various χ approximations. J Royal Stat Soc: Series B (Methodological) 2018;16(2):296-298. [doi: 10.1111/j.2517-6161.1954.tb00174.x]
38. Norris M, Lecavalier L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. J Autism Dev Disord 2010;40(1):8-20. [doi: 10.1007/s10803-009-0816-2] [Medline: 19609833]
39. Roberson RB, Elliott TR, Chang JE, Hill JN. Exploratory factor analysis in rehabilitation psychology: a content analysis. Rehabil Psychol 2014;59(4):429-438. [doi: 10.1037/a0037899] [Medline: 25221958]
40. Dinno A. Implementing horn's parallel analysis for principal component analysis and factor analysis. The Stata Journal: Promoting communications on statistics and Stata 2009;9(2):291-298. [doi: 10.1177/1536867x0900900207]
41. Schweizer K, DiStefano C. editors. Principles and Methods of Test Construction: Standards and Recent Advances. Boston, MA: Hogrefe; 2016.
42. Fabrigar LR, Wegener DT. Exploratory Factor Analysis. Oxford, United Kingdom: Oxford University Press; 2011.
43. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107(2):238-246 [FREE Full text] [doi: 10.1037/0033-2909.107.2.238] [Medline: 2320703]
44. Yuan KH, Chan WS, Marcoulides GA, Bentler PM. Assessing structural equation models by equivalence testing with adjusted fit indexes. Structural Equation Modeling: A Multidisciplinary Journal 2015;23(3):319-330. [doi: 10.1080/10705511.2015.1065414]
45. Hooper D, Coughlan J, Mullen M. Structural equation modeling: guidelines for determining model fit. EJBRM 2007:6 [FREE Full text]
46. Hardouin JB, Bonnaud-Antignac A, Sébille V. Nonparametric item response theory using stata. The Stata Journal: Promoting communications on statistics and Stata 2011;11(1):30-51. [doi: 10.1177/1536867x1101100102]
47. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A practical guide to their development and use. Oxford, United Kingdom: Oxford University Press; 2014.
48. Flora DB, Labrish C, Chalmers RP. Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. Front Psychol 2012;3:55 [FREE Full text] [doi: 10.3389/fpsyg.2012.00055] [Medline: 22403561]
49. Price LR. Psychometric methods: Theory into practice. New York, NY: Guilford Press; 2017.
50. Gaskin CJ, Happell B. On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. Int J Nurs Stud 2014;51(3):511-521. [doi: 10.1016/j.ijnurstu.2013.10.005] [Medline: 24183474]

51. Meehl PE. Why summaries of research on psychological theories are often uninterpretable. PR 1990;66(1):195. [doi: 10.2466/pr0.66.1.195-244]
52. Liu Y, Lu X, Zhao G, Li C, Shi J. Adoption of mobile health services using the unified theory of acceptance and use of technology model: self-efficacy and privacy concerns. Front Psychol 2022;13:944976 [FREE Full text] [doi: 10.3389/fpsyg.2022.944976] [Medline: 36033004]
53. Chen K, Lou VWQ. Measuring senior technology acceptance: development of a brief, 14-Item scale. Innov Aging 2020;4(3):igaa016 [FREE Full text] [doi: 10.1093/geroni/igaa016] [Medline: 32617418]
54. AshaRani PV, Lai D, Koh J, Subramaniam M. Purpose in life in older adults: a systematic review on conceptualization, measures, and determinants. Int J Environ Res Public Health 2022;19(10):5860 [FREE Full text] [doi: 10.3390/ijerph19105860] [Medline: 35627396]
55. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Applied Sciences 2021;11(22):10537. [doi: 10.3390/app112210537]
56. Alsyouf A, Lutfi A, Alsubahi N, Alhazmi FN, Al-Mugheed K, Anshasi RJ, et al. The use of a technology acceptance model (TAM) to predict patients' usage of a personal health record system: the role of security, privacy, and usability. Int J Environ Res Public Health 2023;20(2):1347 [FREE Full text] [doi: 10.3390/ijerph20021347] [Medline: 36674105]
57. Hsiao CH, Tang KY. Examining a model of mobile healthcare technology acceptance by the elderly in taiwan. Journal of Global Information Technology Management 2015;18(4):292-311. [doi: 10.1080/1097198x.2015.1108099]
58. Nadal C, Sas C, Doherty G. Technology acceptance in mobile health: scoping review of definitions, models, and measurement. J Med Internet Res 2020;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
59. Yau Y, Hsiao CH. The technology acceptance model and older adults' exercise intentions-a systematic literature review. Geriatrics (Basel) 2022;7(6):124 [FREE Full text] [doi: 10.3390/geriatrics7060124] [Medline: 36412613]
60. Noorbergen TJ, Adam MTP, Teubner T, Collins CE. Using co-design in mobile health system development: a qualitative study with experts in co-design and mobile health system development. JMIR Mhealth Uhealth 2021;9(11):e27896. [doi: 10.2196/27896] [Medline: 34757323]
61. Taherdoost H. Importance of technology acceptance assessment for successful implementation and development of new technologies. GJES 2019;1(3). [doi: 10.33552/gjes.2019.01.000511]
62. QuillBot. URL: https://quillbot.com/ [accessed 2025-03-03]

Abbreviations

aOR: adjusted odds ratio
AUROC: area under the receiver operating characteristic
CFA: confirmatory factor analysis
CFI: comparative fit index
COSMIN: Consensus-Based Standards for the Selection of Health Status Measurement Instruments
EFA: exploratory factor analysis
FC: facilitating condition
FITT: Fit between Individuals, Tasks, and Technology
Health-ITUES: Health Information Technology Usability Evaluation Scale
IRT: item response theory
ITC: International Test Commission
KMO: Kaiser-Meyer-Olkin
MAUQ: Mobile Health App Usability Questionnaire
mHealth: mobile health
PCA: principal component analysis
R&D: research and development
REDCap: Research Electronic Data Capture
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
STAM: senior technology acceptance model
SUS: System Usability Scale
TAM: technology acceptance model
TLI: Tucker-Lewis Index
UTAUT: unified theory of acceptance and use of technology

Edited by Y Jiang; submitted 03.05.24; peer-reviewed by S Hejazi, M Lee; comments to author 02.12.24; revised version received 17.12.24; accepted 19.02.25; published 11.03.25.

Please cite as:
Buawangpong N, Siviroj P, Pinyopornpanish K, Sirikul W
Transcultural Adaptation, Validation, Psychometric Analysis, and Interpretation of the 22-Item Thai Senior Technology Acceptance Model for Mobile Health Apps: Cross-Sectional Study
JMIR Aging 2025;8:e60156
URL: https://aging.jmir.org/2025/1/e60156
doi:10.2196/60156
PMID:

Original Paper

Developing a Dyadic Immersive Virtual Environment Technology Intervention for Persons Living With Dementia and Their Caregivers: Multiphasic User-Centered Design Study

Elizabeth A Rochon[1], BA; Ayush Thacker[2], BS; Mirelle Phillips[3], BA; Christine Ritchie[2,4*], MSPH, MD; Ana-Maria Vranceanu[1,4*], PhD; Evan Plys[1,4], PhD

[1]Center for Health Outcomes and Interdisciplinary Research, Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

[2]Mongan Institute Center for Aging and Serious Illness and the Division of Palliative Care and Geriatric Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[3]Studio Elsewhere, Brooklyn, NY, United States

[4]Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Evan Plys, PhD
Center for Health Outcomes and Interdisciplinary Research
Department of Psychiatry
Massachusetts General Hospital
1 Bowdoin Square
Boston, MA, 02114
United States
Phone: 1 617 724 0200
Email: eplys@mgh.harvard.edu

Abstract

Background: Persons living with dementia and their caregivers experience frequent emotional health challenges. Across the illness spectrum, engaging in shared pleasant activities is an important feature of well-being for persons living with dementia–caregiver dyads. Under the umbrella of virtual reality, immersive virtual environment technology (IVET) offers artificial sensory experiences and shows promise in this population. IVET development benefits from a user-centered design approach, and as an emerging field, preliminary testing of safety, usability, and engagement for person living with dementia–caregiver dyads is required.

Objective: We aimed to develop a preliminary IVET intervention for psychosocial health among person living with dementia–caregiver dyads. In doing so, we highlight design considerations and user preferences to ensure the safety and usability of technology-based interventions in the context of dementia.

Methods: We engaged 10 clinicians, 8 caregivers, and 3 persons living with dementia in 5 rounds of focus groups to evaluate the safety and usability of preliminary intervention features. Following prototype development, we engaged caregivers and persons living with dementia (n=9 dyads) in beta testing workshops to observe real-time user interaction with the intervention and guide refinements. Rapid data analysis was used to extract themes relevant to intervention development.

Results: The following themes emerged from focus groups to inform prototype development: (1) designing flexibly to allow users to tailor the intervention experience to their own environmental context and circumstance, (2) designing with the dyad's clinical and relational needs in mind, and (3) accounting for illness and aging-related challenges in design. The following themes emerged from workshops to inform prototype refinements: (1) increasing user support through more feedback and (2) increasing variety of visual and auditory feedback.

Conclusions: Using user feedback throughout the development process, we developed a prototype of an IVET intervention, Toolkit for Experiential Well-Being in Dementia (the Isle of TEND), tailored to the needs of persons living with dementia and their caregivers. Our prototype uses specific design features to promote safety, usability, and engagement in the context of dementia. Future feasibility testing of the intervention is warranted.

International Registered Report Identifier (IRRID): RR2-10.2196/52799

(JMIR Aging 2025;8:e66212) doi:10.2196/66212

KEYWORDS

dementia; caregivers; dyads; technology; design; immersive

Introduction

Background

Throughout the course of dementia, emotional health challenges (eg, depression, negative affect, and worry) are common among both persons living with a dementia diagnosis and their informal caregivers (eg, family members and friends) [1,2]. Persons living with dementia and their caregivers present with different emotional needs throughout the course of dementia [3,4]; however, maintaining relationship quality within the context of changing interpersonal dynamics remains an important component of emotional health for both members of the person living with dementia–caregiver dyad [5].

Dyadic research is gaining popularity within the fields of emotional health and dementia care [6,7]. Dyadic research highlights the impact of interpersonal relationships and the transmission of emotional states between members of the person living with dementia–caregiver dyad, which influence the emotional health of each member [8]. For example, perceived relationship strain by the persons living with dementia is associated with increased emotional distress in their caregivers and vice versa [9,10]. Alternatively, positive affective states can be shared among dyad members (eg, laughter), increasing emotional health for both dyad members [11]. Indeed, relationships and the emotional health needs of persons living with dementia and caregivers often change throughout the course of dementia, especially in advanced stages, as cognitive and communication limitations are more common [4,12]. However, despite changes, the need for positive experiences shared between persons living with dementia and caregivers remains common and relates to fewer behavioral and psychiatric symptoms among persons living with dementia and less stress among caregivers [13]. Thus, shared pleasant activities are an important component of emotional health among person living with dementia–caregiver dyads that can be targeted at any stage of illness.

As dementia advances and cognitive impairments become more pronounced, identifying shared activities that can promote emotional health for both persons living with dementia and caregivers may become more challenging [14-16]. Indeed, many dyadic interventions in the context of dementia are concerned with systematic approaches to identifying and engaging in shared activities [17]. For example, the tailored activity program intervention helps dyads identify and engage in personalized shared activities to reduce agitation among persons living with dementia and improve well-being among caregivers [18,19]. However, many of these shared activity interventions rely heavily on caregivers to plan and initiate activities, which may reduce shared positive affective states due to the caregiver-care recipient role, rather than equitable activity engagement [11]. For example, caregivers may find themselves in the role of an “instructor” or “supervisor,” rather than fully engaging in and receiving the emotional benefits of an activity. Therefore, designing interventions that allow both persons living with dementia and caregivers to fully engage in shared activities may create a positive emotional experience for each dyad member, contributing to their emotional health. Technology may aid this endeavor by creating interactive experiences that minimize the need for caregiver initiation (ie, using a technology platform that engages and guides both members of the dyad through an activity, rather than relying on the caregiver to lead the activity). Immersive virtual environment technology (IVET) is one such option for engaging person living with dementia–caregiver dyads in an intervention together.

IVET exists under the umbrella of virtual and augmented reality [20] and is defined as an artificial sensory experience that users actively engage with as if it were a real experience [21]. IVET promotes active engagement between users and computer-simulated environments through multisensory user-driven experiences [13,21,22]. These methods align well with existing approaches to promoting positive affect among persons living with dementia (ie, multisensory stimulation and personalized activities) [23]. IVETs often follow the 3 I’s design framework [24], which suggests that user engagement is achieved through (1) immersion (ie, creating a feeling of presence within the virtual environment), (2) interaction (ie, creating a sense of agency and control over the virtual environment), and (3) imagination (ie, the potential to create new experiences within the virtual environment). In nondementia populations, IVET interventions have demonstrated positive effects on emotional health outcomes, such as positive affect and emotional health (eg, depression and stress) [25]. Importantly, IVET interventions vary in terms of the level of immersion, categorized as low (ie, screen-based intervention), medium (ie, cave automatic virtual environments and surrounding projection screens), or high (ie, head-mounted displays for isolation from physical reality) [25]. In nondementia populations, medium immersion produces the greatest results on emotional health, with little difference observed between low and high immersion [25]. However, given the potential safety risks, such as cybersickness, falls, and agitation associated with higher levels of immersion in the context of dementia [21,26], IVET with low immersion may be suitable for this population. Indeed, most IVET interventions tested within the context of dementia use lower levels of immersion [27].

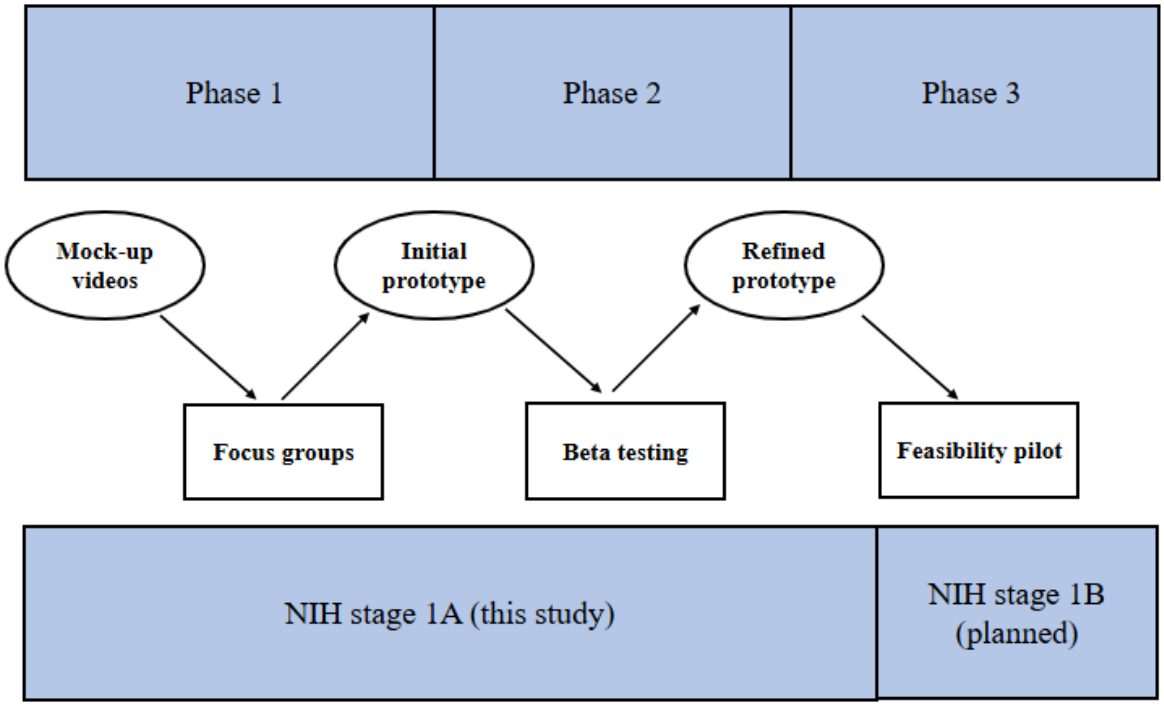
IVET interventions allow for multiple users to engage jointly (eg, persons living with dementia and their caregivers participate in shared activities) [21]. However, existing interventions have been tested primarily among persons living with dementia and caregivers individually [27-30]. For persons living with dementia, IVET interventions have demonstrated feasibility and usability, primarily targeting cognitive rehabilitation, exercise, or spatial navigation training [27]. For caregivers, IVET interventions have demonstrated feasibility and usability, often focused on caregiver education and skills [31,32]. Fewer IVET interventions have been developed for person living with

dementia–caregiver dyads to participate in jointly (ie, dyadically), with a focus on shared positive experiences for emotional health. SENSE-GARDEN is an example of a dyadic IVET intervention, which creates an immersive environment based on preferred autobiographical experiences, such as music, images, videos, and personal media (eg, family photos), for persons living with dementia who have moderate to severe cognitive limitations, allowing them to navigate along with their caregivers. The intervention is designed to support reminiscence between the persons living with dementia and caregivers in the present moment. Notably, caregivers have an active role in personalizing the information in SENSE-GARDEN but receive a lower "active dose" of the intervention (ie, fewer caregiver-focused experiences). Rather, caregiver benefit is conceptualized as arising from positive interactions with persons living with dementia. A clinical trial is currently underway for this intervention [33].

A major challenge to designing dyadic IVET interventions in the context of dementia is addressing cognitively mismatched dyads, which is more common as dementia advances. Promoting shared engagement among person living with dementia–caregiver dyads across the illness spectrum must embrace the 3 I's of design (ie, immersion, interaction, and imagination) for dyads with potentially different cognitive ability levels. While challenging from a design perspective, if achieved, this type of intervention would be an invaluable opportunity for dyad members to truly engage in shared activities together (ie, not initiated by the caregiver), helping them to briefly abandon the caregiver-care recipient dynamic and, in doing so, derive relational and emotional benefits. These types of more equitable interactions may promote positive affect and interpersonal connection among both persons living with dementia and caregivers, dignity for the persons living with dementia, and positive aspects of caregiving for the caregivers [34]. However, currently, guidelines for designing dyadic IVET interventions to promote shared engagement among person living with dementia–caregiver dyads are lacking, especially for dyads with mismatched cognitive abilities [6,35].

Additional work is needed to guide the development of dyadic IVET interventions throughout the course of dementia, especially in more advanced stages. For this endeavor, user-centered design approaches may be particularly important to ensure that interventions have the potential to be acceptable and feasible for both dyad members at various stages of illness. User-centered design is an iterative approach that centers the user experience early and often throughout the intervention development process [36]. The aim of user-centered design is to maximize the feasibility and acceptability of an intervention [37]. This study followed a multiphasic conceptual model of user-centered design rooted in human factors research, as outlined by Witteman et al [38]; refer to Figure 1. In phase 1, design teams engaged users and other knowledge holders with the goal of understanding user needs, goals, and strengths. In phase 2, design teams iteratively developed and refined an initial intervention prototype with user feedback. In phase 3, design teams observed users' naturalistic interactions with the initial prototype to inform additional refinements [38]. On the basis of user-centered design principles in the context of dementia [39], we prioritized the following factors across each design phase: understandability (ie, the user understands the intervention functions), learnability (ie, the user is able to easily learn the intervention), usability (ie, the user is able to participate in a task with minimal assistance), and safety (ie, tasks do not inadvertently lead to negative affect or other adverse outcomes).

Figure 1. User-centered iterative design to develop immersive virtual environment technology. Circles denote activities initiated by the design team, and rectangles denote activities designed to elicit user feedback. NIH: National Institutes of Health.



Purpose

This study aimed to develop a dyadic IVET intervention to support the emotional health of person living with dementia–caregiver dyads. Our design team represents an academic-industry partnership. In this study, we highlight our multiphasic user-centered design process [38], focused on the early stages of intervention development of a dyadic IVET intervention (phases 1 and 2 of user-centered design). First, we present data from focus groups with clinicians, caregivers, and person living with dementia–caregiver dyads to understand user needs, goals, strengths, and usability and safety concerns (phase 1). Then, we discuss the process of iteratively developing and refining an initial intervention prototype based on user feedback from multiple user workshops with person living with dementia–caregiver dyads (phase 2). The resulting dyadic IVET intervention prototype will be ready for pilot testing in a pilot feasibility trial (phase 3). In this study, by describing phases 1 and 2 of user-centered design, we highlight early design considerations for future teams focused on developing IVET or other novel technologies with person living with dementia–caregiver dyads across stages of dementia.

Methods

Study Design

This study followed a multiphasic user-centered design framework for the early stages of intervention development, as outlined by Witteman et al [38] and with the following dementia-specific considerations incorporated [39]: (1) understanding user needs, goals, strengths, limitations, and safety concerns; (2) iteratively developing and refining an initial prototype with user feedback; and (3) observing users' naturalistic interactions with the initial prototype. In phase 1, we conducted 5 focus groups with clinicians, caregivers, and person living with dementia–caregiver dyads to identify preferred activity features as well as design considerations for usability (ie, preferred intervention activities and accommodations) and safety in the context of dementia. The study team determined readiness to move to phase 2 after we believed that we understood major safety risks and could generate enough content for mock-ups of an intervention prototype. In phase 2, we conducted 8 beta testing workshops with person living with dementia–caregiver dyads to inform the development and refinement of an initial prototype. We determined readiness to move to phase 3 when usability concerns were sufficiently addressed, as evidenced by minimal usability issues reported by workshop participants. At the conclusion of phase 2, we had a finalized intervention prototype ready for feasibility testing in phase 3, which is not reported in this study. Phases 1 and 2 map onto stage 1A of the National Institutes of Health (NIH) Stage Model of Behavioral Intervention Development (ie, intervention generation, refinement, modification, and adaptation), and phase 3 maps onto stage 1B (ie, pilot testing) [40,41].

Phases 1 and 2 of intervention development were informed by our intervention conceptual model, as suggested by Rochon et al [42]. Specifically, we aimed to develop an IVET intervention that could promote sustained attention, positive emotions, and active engagement (ie, mechanistic targets) among both persons living with dementia and caregivers, leading to increased relationship satisfaction, decreased distress, and reduced agitation among persons living with dementia, as well as decreased burden among caregivers (ie, outcomes).

Ethical Considerations

Procedures were approved by the Massachusetts General Hospital Institutional Review Board (2022P001401). The information reported in this manuscript follows the Guidance for the Reporting of Intervention Development standards [43]. Potential clinician participants provided verbal consent before participating in focus group interviews. Caregivers provided verbal consent or e-consent during screening phone calls. A research assistant administered the University of California, San Diego Brief Assessment of Capacity to Consent Questionnaire (UBACC [44]) to persons living with dementia to determine their ability to provide consent. For those unable to provide consent, the persons living with dementia provided assent, and their caregivers provided surrogate consent to participate. All data were deidentified to protect anonymity. In phase 1, clinicians were compensated US $30, and persons living with dementia and caregivers were compensated US $50 each. In phase 2, persons living with dementia and caregivers were compensated US $50 each for their time.

Participant Recruitment

During phases 1 and 2, this study enrolled three categories of participants: (1) clinicians involved in dementia care, (2) caregivers of persons living with dementia, and (3) person living with dementia–caregiver dyads.

Clinicians were recruited via emails to dementia care clinics within the affiliated health care system. Clinicians were eligible if they (1) self-identified as dementia care providers and (2) were able to participate via Zoom (Zoom Communications). Of the 11 clinicians screened for focus groups, 1 (9%) did not participate due to a scheduling conflict.

Caregivers and dyads were both recruited via research registries (eg, study fliers). Recruitment materials were designed to target caregivers. Interested caregivers responded to the research team to obtain additional information about the study and review eligibility. Caregivers were eligible if they (1) aged >18 years, (2) self-identified as informal caregivers for a person living with dementia (eg, a family member or friend providing support), (3) could participate in English, and (4) had access to a video-enabled device for Zoom participation. Of the 31 caregivers screened for focus groups, 5 (16%) were ineligible due to not identifying as a caregiver of a person living with dementia, 13 (42%) had scheduling conflicts, and 5 (16%) could not be contacted.

For person living with dementia–caregiver dyads, caregivers were targeted using similar recruitment strategies discussed above. Interested caregivers identified and contacted their care recipient with dementia to participate in the study. Then, caregivers and persons living with dementia participated in a screening call together. Using the same eligibility criteria stated in the Participant Recruitment section, caregivers provided verbal consent or e-consent during screening calls. Eligibility

criteria for persons living with dementia included (1) having an informal caregiver participating in the study, (2) a self- or caregiver-reported dementia diagnosis, and (3) the ability to participate in an interview conducted in English. The UBACC uses 2 open-ended questions to assess the understanding of informed consent information (eg, "Can you briefly tell me what the purpose of the study is based on what I just read?"); scores range from 0 to 2, with scores <2 indicating lack of capacity to provide consent. The UBACC scores show a strong correlation with global cognition [45], suggesting that lower scores (ie, the inability to provide consent) are indicative of more advanced stages of dementia. Across stages, of the 12 persons living with dementia screened for this study, 8 (67%) scored a 0, a total of 2 (17%) scored 1, and 2 (17%) scored 2 (ie, able to consent independently). Of the 26 caregivers screened for workshops, 1 (4%) was ineligible due to not identifying as a caregiver, 4 (15%) had scheduling conflicts, 3 (12%) declined due to the care partner's worsening medical status, and 9 (35%) could not be contacted.

Phase 1: Focus Groups

Participants

There were 21 participants across 5 focus groups. A total of 10 clinicians participated across 2 clinician-only focus groups, representing the following professions: geriatric medicine (n=2, 20%), neurology (n=3, 30%), social work (n=2, 20%), occupational therapy (n=1, 10%), psychology (n=1, 10%), and psychiatry (n=1, 10%). Clinician degrees included Occupational Therapist (1/10, 10%), Licensed Psychologist (1/10, 10%), Master of Social Work (1/10, 10%), Licensed Clinical Social Worker (1/10, 10%), Licensed Independent Clinical Social Worker (1/10, 10%), Doctor of Medicine (6/10, 60%), Doctor of Science (1/10, 10%), and Doctor of Philosophy (PhD; 3/10, 30%). Participants reported an average of 10.95 (SD 10.20; range 2-20) years of experience in dementia care. The majority were women (6/10, 60%) and White (9/10, 90%) individuals.

A total of 4 caregivers participated in a caregiver-only group. Participants were primarily women (4/4, 100%), and half identified as White (2/4, 50%) individuals.

A total of 3 person living with dementia–caregiver dyads (n=6) and 1 solo caregiver (ie, the corresponding person living with dementia was unable to attend) participated across 2 dyad focus groups. Persons living with dementia were primarily men (2/3, 67%) and White (2/3, 67%) individuals. Caregivers were primarily women (4/4, 100%), and half identified as White (2/2, 50%) individuals.

Procedure

The sequence of focus groups was clinicians (n=2), person living with dementia–caregiver dyads (n=2), and caregiver-only (n=1) group. A PhD-level clinical psychologist with experience in dementia care and qualitative research led 4 focus groups with clinicians and person living with dementia–caregiver dyads. During dyadic focus groups, the facilitator made every effort to ensure active participation from persons living with dementia, including seeking collateral information from dyad members (eg, confirming caregiver statements with persons living with dementia), asking binary questions (eg, yes or no), rephrasing questions when needed, and making behavioral observations with verbal confirmations (eg, "I noticed you perked up when I mentioned music, do you enjoy listening to music?"). Two trained bachelor-level research assistants coled the caregiver-only group. None of the interviewers had previous relationships with the participants. Group sessions lasted an average of 54 (range 51–57) minutes. All focus groups were completed virtually over secure Zoom, audio recorded, and transcribed.

During each focus group, participants responded to semistructured interview questions targeting activity preferences, technology preferences, usability considerations, and safety concerns; refer to Multimedia Appendix 1 for example focus group questions. In addition, participants provided feedback on a series of mock-up videos (ie, disjointed visuals designed to elicit conceptual feedback and ideas for intervention modules), including passive nature views (ie, waves crashing and Northern lights), guided relaxation with nature background (ie, night sky with voice-over), and abstract color flow (ie, colorful viscous swirl). Mock-up videos were previously used by the design team to inform the development of other IVET platforms targeting nondementia populations. Videos were chosen to provide insight into design features that aligned with our proposed mechanistic targets of sustained attention, positive emotions, and active engagement [42]. Participant feedback was designed to elicit preferences, usability considerations, and safety concerns (eg, "What problems do you see with this experience?" and "Do you foresee any concerns for safety or otherwise?"). Consistent with an iterative process of design, videos and interview questions were updated between each round of focus groups, such that videos or topics that elicited concerns for usability (eg, querying about the utility of sensory kits of household items, such as spices, to accompany videos, which elicited concern for feasibility) or safety (eg, abstract color flows, which elicited concern about increasing agitation or confusion) were omitted, and others were added (eg, new nature scenes with multiple points of view).

During each focus group, 1 or 2 bachelor-level research assistants observed the interviews and completed live field notes using a standardized rapid data analysis (RDA) template created for use in this study; refer to Multimedia Appendix 1 for the RDA template. At the conclusion of each focus group, all facilitators and observers completed reflective field notes using a standardized form with categories similar to those in the RDA template.

Phase 2: Beta Testing Workshops

Participants

A total of 9 person living with dementia–caregiver dyad (n=18) participants were recruited across 8 beta testing workshops. Dyads were recruited separately for phases 1 and 2; none of the dyads who participated in phase 1 participated in phase 2. However, dyads enrolled in phase 2 were allowed to participate in multiple workshops if desired. Of the 9 dyads, most (6, 67%) participated in 1 workshop and 3 (33%) participated in 2 workshops; however, none participated in >2 workshops. Participants ranged from 1 to 4 dyads per workshop, with the majority consisting of 1 dyad (6/8, 75% workshops).

Across workshops, 100% (18/18) of participants identified as White individuals. Most persons living with dementia (7/9, 78%) identified as male, and most caregivers (7/9, 78%) identified as female. Caregivers provided proxy reports of dementia diagnoses, per their understanding of the medical records of the persons living with dementia, which included Alzheimer disease (2/9, 22%), unspecified dementia (4/9, 44%), Lewy body dementia (1/9, 11%), primary progressive aphasia (1/9, 11%), and frontotemporal dementia (1/9, 11%). Given that this study did not access medical records, caregiver reports were not confirmed, and the stage of illness was not reported. Multiple dyad relationship types were noted, including spousal (4/9, 44%), sibling (1/9, 11%), and adult child (1/9, 11%) caregivers.

Procedure

Beta testing workshops were led by members of the industry partner design team, with 1 to 2 trained research assistants in attendance to assist dyads in navigating the platform and to complete live field notes for RDA. All workshops were completed virtually over secure Zoom and recorded for the research team to review after session completion (eg, detailed feedback and behavioral observations). If multiple dyads engaged in 1 session, breakout rooms were used to enable private viewing experiences. Beta testing workshops lasted an average of 63 (range 34-103) minutes.

Workshops were completed via Zoom. At the start of each workshop, the design team shared a website link to the initial IVET intervention prototype with participants. Dyad participants were instructed to share their screen after opening the website. A workshop facilitator then prompted dyads to navigate the intervention both freely (eg, “As you explore the area, please share your thoughts out loud”) and with specific instructions (eg, “Can you figure out how to navigate back to the home screen?”). Participants started by freely navigating the prototype, providing live feedback on usability and acceptability (eg, confusing instructions and enjoyable features). Then, participants were presented with 1 to 3 modules to navigate, guided by facilitator prompts that primarily focused on usability and acceptability (eg, “Are there specific features on the island that drew your interest?” “What are your initial thoughts on the visuals of the activity?” and “Was navigating around the island intuitive or challenging?”). Dyads could choose the module they were most interested in navigating. In some situations, the second or third module was chosen for the dyad to ensure that feedback was gathered for each module. At the conclusion of each workshop, caregivers and persons living with dementia independently responded to quantitative measures assessing the perceived feasibility and usability of the platform. As workshops progressed, feedback became more targeted (eg, preference for specific images) than in earlier workshops (eg, preferred level of feedback from the platform), moving toward an initial prototype.

During workshops, research assistants completed observational field notes using a standardized template created for use in this study; refer to Multimedia Appendix 1 for the template. Field notes were used as part of RDA to inform intervention refinements between each workshop.

Measures

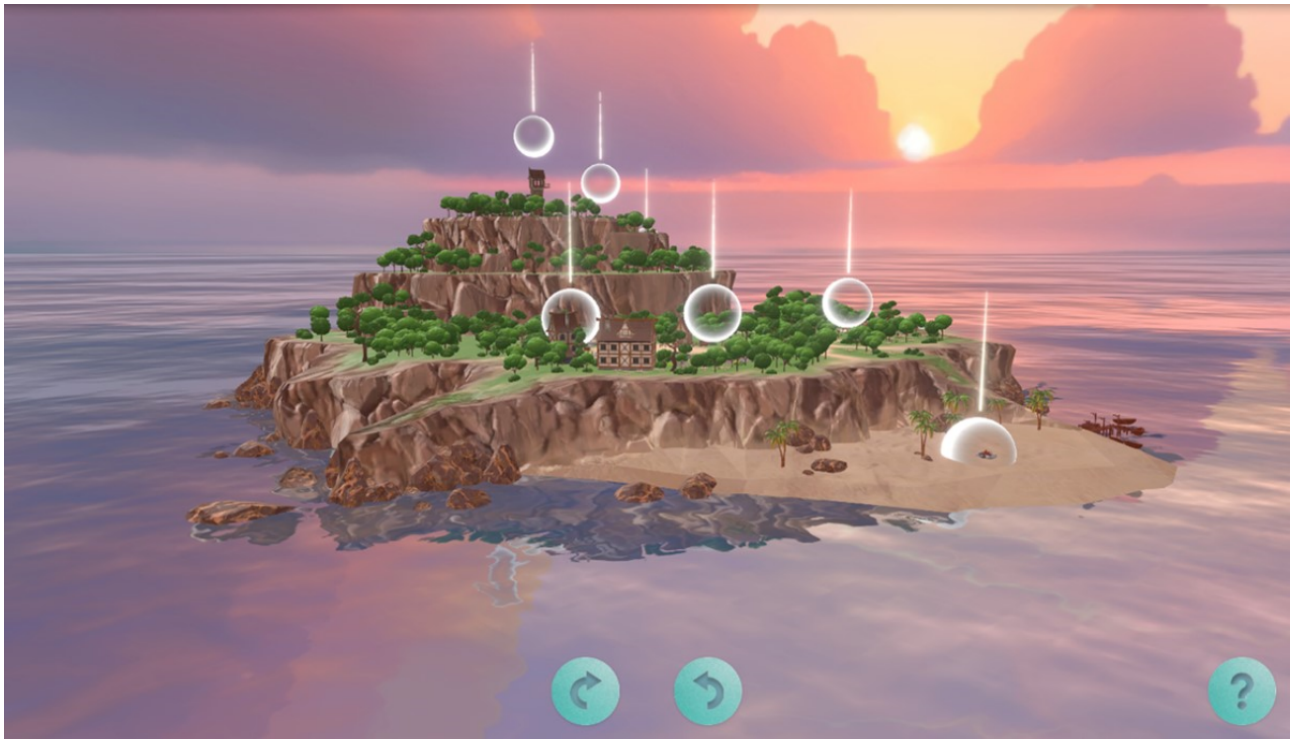
Usability was assessed using the System Usability Scale (SUS) [46]. The SUS consists of 10 items evaluating the perceived usability of technological platforms, including websites. The SUS is the most widely used scale to assess usability for technology-based interventions in the context of dementia, but self-report challenges may exist with more advanced cognitive limitations [47]. However, the SUS is primarily applied to persons living with dementia with mild cognitive impairments and lacks psychometric data across the full spectrum of dementia (ie, limited data available in advanced stages of dementia) [47]. Furthermore, the SUS uses 5-point Likert scale for response options [48,49]. Therefore, following best practices to reduce cognitive load and ensure that persons living with dementia could meaningfully participate, we adapted this scale to reflect yes or no response categories [47,50]. This decision increased the likelihood that persons living with dementia in this study could provide feedback independently.

A total of 3 additional study-specific feasibility questions targeted entertainment, safety, and control. These items were adapted from previous technology-based intervention studies with participants with mild cognitive impairment and their caregivers [48] and provided feedback salient to our conceptual model (eg, engagement and positive affect) and design aims (eg, safety). These items were also asked in a yes or no format to increase inclusivity for persons living with dementia.

Intervention Prototype

Activities outlined in Phase 1: Focus Groups and Phase 2: Beta Testing Workshops sections culminated in the development of an initial prototype of a web-based dyadic IVET intervention, referred to as the Toolkit for Experiential Well-Being in Dementia (Isle of TEND). The Isle of TEND website contained a home page (Figure 2) with the ability to navigate to 8 modules, which contained different shared activities; refer to Multimedia Appendix 2 for an overview of modules. Consistent with our conceptual model [42], each module in the Isle of TEND featured a shared activity intended to promote sustained attention, positive emotions, and active engagement with the intervention. The intervention was designed to be used by both persons living with dementia and caregivers together in the home.

Figure 2. The Toolkit for Experiential Well-Being in Dementia (Isle of TEND) home screen.



Data Analysis

This study used RDA as an analytic framework to identify themes from focus groups and workshops. RDA is used to quickly identify themes in qualitative data [51]. Rather than written transcriptions, this technique relies on live coding directly from interviews, such as taking observational notes during interviews or while listening to an audio recording [52-54]. For this study, RDA was chosen to identify specific feedback from participants related to preferences, usability, and safety that could guide prototype development and refinement. By quickly capturing this information, our research team was able to communicate findings back to our design team to iteratively respond to feedback and update the intervention prototype to gather more feedback from users.

Consistent with previous studies using RDA [55], we compiled RDA template notes from 2 to 3 team members (ie, 1 PhD-level clinical psychologist and 2 bachelor-level research assistants) into a single matrix for each focus group. Data from this matrix were used to guide discussions with the design team and informed design decisions toward initial prototype development.

At the conclusion of the study, matrices were combined and used to facilitate thematic analysis of focus groups and workshops separately. Focus group data were analyzed as a single dataset (ie, clinician, caregiver, and dyad focus groups analyzed together), such that we identified themes that emerged across groups. We did not compare themes across populations, as this was not consistent with the aims of this study. Two research assistants (ER and AT) trained in qualitative RDA methods independently reviewed matrices to identify codes. Coders met on multiple occasions during the coding process to resolve discrepancies. Codes were grouped by content to form themes. The thematic structure was iteratively developed and confirmed with a meeting between coders and a clinical psychologist on the study team with experience in dementia care (EP). After a thematic structure was agreed upon, the research assistants (ER and AT) reviewed transcripts to identify salient quotes for each thematic category. Both focus groups and workshops followed the same analytic steps. However, transcripts were only reviewed for focus groups because the data collected during workshops produced fewer exemplar quotes (eg, workshops contained more “thinking aloud” methods or behavioral observation, making data less relevant for exemplar quotations).

Results

Phase 1: Focus Groups

Three themes emerged from focus groups with clinicians, caregivers, and dyads: (1) designing flexibly to allow users to tailor the intervention experience to their own environmental context and circumstance, (2) designing with the dyad’s clinical and relational needs in mind, and (3) accounting for illness and aging-related challenges in the design.

Theme 1: Designing Flexibly to Allow Users to Tailor the Intervention Experience to Their Own Environmental Context and Circumstance

The ability for users to individualize their experience emerged as an important design feature to increase acceptability and usability as well as to empower choices among users. Participants noted that the option to individualize or tailor intervention content could help accommodate a larger range of users based on their access to and comfort navigating technology. In addition, participants recommended that the IVET intervention be accessible via various platforms. For example, a clinician noted the following:

> *[Consider] an app that is free, that can download the materials into [users] cell phone or their device so it doesn't depend on the Internet connection...and maybe having the option of audio only or audio plus visual.*

Participants also noted the importance of developing stimuli that were familiar to and pleasant for participants of various backgrounds, cohorts, and identities to promote individualization. Furthermore, participants noted that familiar stimuli may especially increase comfort, enjoyment, engagement, and safety for persons living with dementia. A caregiver highlighted the potential benefits of familiar stimuli:

> *Having an inspirational waterfall isn't necessarily going to put me at ease, but maybe hearing my family's familiar voice, or the clattering of my own kitchen, or having something that individualizes this to me would.*

Similarly, a clinician highlighted the need for inclusivity as follows:

> *I just wonder if the salience of some of the images might be different depending on the cultural groups. The aurora borealis is beautiful, but for maybe people who [are] not familiar with that part of the of the world...what are the scenes, the sounds that are most salient and comforting to that individual.*

Of note, related to the environmental context of an at-home IVET intervention, caregivers discouraged the use of a proposed sensory kit of household materials, citing feasibility issues with the ability for a range of users to locate these materials in their home (eg, spices).

Theme 2: Designing With the Dyad's Clinical and Relational Needs in Mind

Participants noted that designing an intervention with a specific purpose and orienting users to a rationale for using the intervention may increase usability. Specifically, participants suggested recommendations for appropriate times and intended outcomes of use (eg, using when calm and avoiding when agitated). For example, a clinician noted the potential clinical utility of an IVET intervention:

> *I think that [the calming sounds are] good, and it appeals to an inherent positive emotional experience which could also aid caregivers when they're trying to prep for redirection or transition. Basically, if we're trying to get somebody to do something that they may not necessarily want to do we're far more likely to be able to if a person is in a positive emotional state versus a negative one and so these types of things could aid in that process of prepping a patient for a transition.*

Caregivers believed that they would likely need to be present to initiate the intervention as well as to guide the persons living with dementia through some of the modules. For example, a caregiver noted the following:

> *[Caregivers] would have to be there to sort of guide [Person living with dementia] or kinda help them figure out what [the intervention] is.*

Participants were skeptical that an IVET intervention could be engaging for both dyad members, noting challenges of achieving sustained attention and immersion for both persons living with dementia and caregivers. This is highlighted by a caregiver who stated the following:

> *[Person living with dementia] loses interest pretty quick, and I think sometimes he has trouble following the storyline. So, it would be hard to find something that would be suitable for both of us, cause we're kind of in very different places now. And that's one of the things that's difficult is we don't have as many shared experiences as we used to.*

Participants recommended structured stimuli to quickly capture and sustain the attention of users, which they noted could be challenging in the context of dementia. For example, a caregiver stated the following:

> *Without having a clear beginning or end, [the person living with dementia] might get bored or discouraged, it needs to have more structure.*

Furthermore, participants emphasized that the IVET intervention would need to be designed to sustain attention and immersion without overstimulating or confusing users with cognitive challenges.

Theme 3: Accounting for Illness and Aging-Related Challenges in the Design

Participants noted the importance of ensuring the reduction, rather than an unintentional exacerbation, of behavioral and psychiatric symptoms of dementia, particularly agitation and hallucinations. This concern is highlighted by a clinician, who stated the following:

> *Dementias with any psychiatric component or psychotic features, paranoia, delusional thinking...you always want to be careful about providing stimulation when we don't know where that may lead the patient's train of thought.*

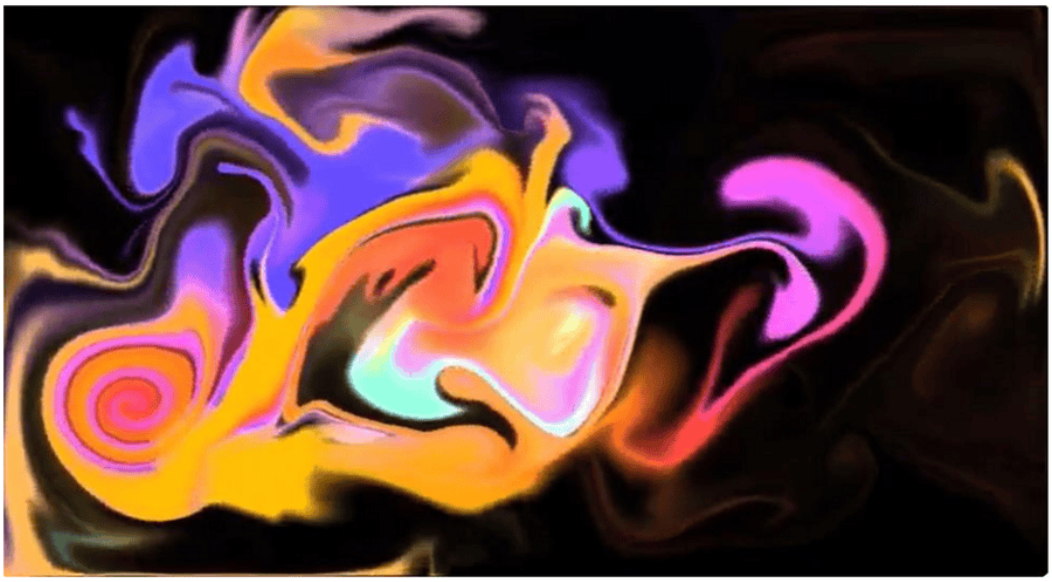
There were specific stimuli that participants recommended avoiding, such as 1 video with a moving (ie, swirling) bright color scheme against a black background (Figure 3), which participants suggested could potentially exacerbate psychiatric symptoms. For example, a clinician stated the following:

> *I think patients would not like [the swirling color video], they have a lot of visual spatial misperceptions.*

Furthermore, participants cautioned against dark backgrounds, which might increase psychiatric symptoms and confusion or disorientation. For example, a caregiver noted the following:

> *My husband is really paranoid, he thinks people are out to get him. I was thinking of the lighting, if there is darkness that might increase his paranoia.*

Figure 3. Swirling color scheme video.



Another caregiver noted the following:

> *The edges of my video were very dark. So, it gives you almost like that movie theater feel to it, which I worry, if you're immersed in something too much, do you have less grasp on here and now where you are.*

In addition, as noted across multiple themes, participants relayed that certain cognitive symptoms, especially attentional deficits and apathy, could pose a challenge to engagement and usability. One clinician stated the following:

> *I wonder if in some instances this [mock-up videos] may be a little too unstimulating to capture the patient's attention, unless they are really already very calm and wanting to be calm.*

Another clinician stated the following:

> *for images...it might be really, really simple ones, the less complicated you make it, the better received because [person living with dementia] cognitive abilities and processual abilities are impaired.*

Participants also noted that visual clarity was an important design feature and recommended that the IVET intervention accommodate for age-related changes and pathologies more common with older age. For example, 1 caregiver stated the following:

> *There's something about lighting that I would pay attention to for elders...I have another elder in my life who has glaucoma and she can't see things in depth of perception and other things.*

Phase 2: Beta Testing Workshops

Qualitative Results

Two main themes emerged from beta testing workshops with dyads that informed intervention refinement: (1) more feedback to guide using the intervention and (2) more variety of visual and auditory experiences.

Theme 1: Increasing User Support Through More Feedback

Participants noted that more purposeful or goal-directed actions within intervention modules could reduce confusion during use. When engaging with the intervention, caregivers and persons living with dementia frequently inquired about the end goal of modules, especially when feeling "lost" or "stuck." Relatedly, participants recommended more instructions, which, per observation, appeared to be related to participant concerns about making an error or doing the experience "wrong." Together, an open-ended module without clear directions, which was initially developed to increase exploration and reduce the need for caregiver instructions, actually increased confusion and perceived misuse of the intervention by users.

Furthermore, participants voiced the need for continuous instructions throughout use to understand next steps, which was particularly relevant in modules that included multistep tasks. For example, a module that developed a poem based on typed responses to multiple prompts caused some confusion among users because dyads were unsure of the task they were being prompted to complete (eg, not knowing that prompts would generate a poem) or the end goal of the multistep process (eg, the need to press next and respond to another prompt to develop the full poem). In addition to more prompts, participants suggested having multiple methods of engagement, such as a speech-to-text function, to promote accessibility and usability.

Theme 2: Increasing the Variety of Visual and Auditory Feedback

Participants recommended increasing the presence and variability of responsive sounds, background music, and interactive visuals. Caregivers and persons living with dementia both noted that alternating music or background colors might reduce redundancy after multiple uses and make the intervention more engaging. In addition to increasing engagement, dyads noted that sound effects, especially calming nature sounds, improved their experience by creating a sense of calm while using the intervention.

Quantitative Analysis

In general, users rated the intervention positively (Table 1). On the study-specific feasibility scale, 100% (11/11) of caregivers and 67% (6/9) of persons living with dementia (n=2, 22% were unsure or did not respond) felt that the program was safe to use, 82% (9/11) of caregivers and 100% (9/9) of persons living with dementia felt that the program was entertaining, and 82% (9/11) of caregivers and 89% (8/9) of persons living with dementia felt that the program helped them "feel in charge of their own experience." On the SUS, 82% (9/11) of caregivers and 89% (8/9) of persons living with dementia thought that the program was "easy to use," and 73% (8/11) of caregivers and 78% (7/9) of persons living with dementia "felt very confident using the program" (n=1, 11% were unsure or did not respond).

Table 1. Feasibility and usability results of beta testing workshops.

	Caregivers, (n=11), n (%)			Persons living with dementia (n=9), n (%)		
	Yes	No	Unsure or unable to answer	Yes	No	Unsure or unable to answer
"I think that I would like to use this program."	7 (64)	2 (18)	2 (18)	6 (67)	2 (22)	1 (11)
"I found the program unnecessarily complex."	1 (9)	10 (91)	0 (0)	1 (11)	8 (89)	0 (0)
"I thought the program was easy to use."	9 (82)	2 (18)	0 (0)	8 (89)	1 (11)	0 (0)
"I think that I would need the support of a technical person to be able to use this program."	4 (36)	7 (64)	0 (0)	3 (33)	4 (44)	2 (22)
"I found the various functions in this program were well integrated."	9 (82)	2 (18)	0 (0)	6 (67)	3 (33)	0 (0)
"I thought there was too much inconsistency in this program."	2 (18)	9 (82)	0 (0)	2 (22)	5 (56)	2 (22)
"I would imagine that most people would learn to use this program very quickly."	7 (64)	3 (27)	1 (9)	4 (44)	0 (0)	5 (56)
"I found the program very cumbersome to use."	0 (0)	11 (100)	0 (0)	1 (11)	5 (56)	3 (33)
"I felt very confident using the program."	8 (73)	2 (18)	1 (9)	7 (78)	1 (11)	1 (11)
"I needed to learn a lot of things before I could get going with this program."	1 (9)	10 (91)	0 (0)	2 (22)	4 (44)	3 (33)
"Was this program able to entertain your relative and you?"	9 (82)	2 (18)	0 (0)	9 (100)	0 (0)	0 (0)
"Did you feel that the program was safe to use for your relative and you?"	11 (100)	0 (0)	0 (0)	6 (67)	1 (11)	2 (22)
"The program helped my relative feel in charge of their own experience."	9 (82)	2 (18)	0 (0)	8 (89)	1 (11)	0 (0)

Feedback-Design Integration

Throughout intervention development, our academic-industry team met weekly to discuss results from focus groups and workshops and translate findings into pragmatic design elements. Often, the research team helped provide the context of findings within the illness context of dementia, as well as guide a systematic approach to intervention design that would prepare the team for a future, scientifically rigorous clinical trial. The industry design team often discussed the possibilities and limitations of an IVET intervention delivered with minimal equipment (ie, via a computer or tablet). These perspectives led to productive discussions about incorporating user feedback into intervention design. Together, this iterative process led to the end product of an IVET intervention ready for initial feasibility testing (NIH stage 1B) that was deemed safe and usable by end users (ie, person living with dementia–caregiver dyads). Feedback and design decisions are provided in Table 2.

Table 2. User feedback and design team responses.

User feedback	Design feature
Focus groups: initial prototyping	
Implement choice-based activities when possible.	The team decided on a web-based IVET[a] format so that users could guide themselves through the platform. The intervention was user driven, and users chose from various modules in the intervention, deciding the type and length of activity within each module.
Establish the ability to adjust audio or visuals depending on preferences.	Developers enabled access to the intervention on the user's own tablet or computer so that users could adjust the volume and brightness to their preference.
Include visuals or audio that are appealing and familiar to diverse populations.	The design team developed 8 different modules, including modules set in a forest, an ocean, a beach, and a lighthouse.
Integrate instructions for the intervention.	The design team included text instructions on the home screen to orient users when they logged into the intervention. The instructions reviewed how to navigate through the home screen and the modules.
Intervention can be used among people with and without cognitive impairment.	Modules were "error-free" so that users were unable to "fail" or choose the "wrong" option, halting the experience; regardless of user choice, the experience continued.
Include stimuli that are reactive to touch or hand or arm movement.	Developers created two activities, which accessed the device's camera to track and react to hand movement: (1) Sound Garden, which used hand movement to control the music pitch and volume in a garden scene, and (2) Art Studio, which used hand movement to create brush strokes on the screen to reveal an art piece.
Design modules that are grounded in reality (familiar structure and clear beginning and ending).	The team set the intervention in recognizable locations, such as a waterfall, a lighthouse, and a beach. The game had a clear beginning with text instructions when users logged on.
Avoid abstract images and darkness in visuals.	The team removed abstract "swirling colors" stimuli. The team discarded a video set in a dark forest environment and brightened the colors in the remaining modules.
Beta testing workshop: prototype refinements	
Clarify the end goal or purpose of the intervention. Include more goal-directed aspects to the intervention.	The design team added an introduction slide clarifying that the intervention was designed to create "shared adventures" between the dyad and "engage the senses and spark imagination." The team included a "help" button across all modules to guide users.
Implement different methods of engagement, such as a speech-to-text option instead of typing.	Developers included a speech-to-text function in modules that required users to type. These activities accessed the device microphone to capture spoken responses from the dyad and convert them into text. To increase engagement, developers created custom soundscapes for scenes and changed them based on the visual stimuli presented.
Add more instructions and audio-to-text instructions.	Developers added a module description when users hovered their mouse over each module location. The team added step-by-step tutorials, which isolated different interactive aspects of the experience when users clicked on each module. In addition, the team added voice-over to text as an accessibility option to all introductions and guides.
Increase the visual stimuli throughout use to reduce redundancy.	The team increased the visual options in modules throughout the experience by creating more artwork and audio clips. For example, in the Storytelling module Campfire Cove, the team increased the number of available pictures to choose from and created an accompanying story.
Include visual and audio stimuli, such as animation or music, in modules that lacked sounds or featured static images. Add calming music or nature audio when possible.	The design team added nature-based animation and sound on the home screen, including auditory and visual waves, so that when logging on, users were greeted with these stimuli. In addition, the design team added nature-based sounds to the background of module instructions, such as birds chirping and "calming" music.

[a]IVET: immersive virtual environment technology.

Discussion

Principal Findings

Following a multiphasic user-centered design framework with dementia-specific considerations [38], this study used focus groups and user workshops to develop and refine the initial prototype of a dyadic IVET intervention to support the emotional health of person living with dementia–caregiver dyads (ie, the Isle of TEND). The intervention was rooted in our conceptual model of shared activities that can promote sustained attention, positive emotions, and active engagement among both dyad members simultaneously [42]. The Isle of TEND is now ready for feasibility testing, consistent with NIH stage 1B.

Design Considerations

Overview

Across focus groups and workshops, we identified several key considerations for designing an IVET intervention for person living with dementia–caregiver dyads. Importantly, we did not constrict our intervention to a specific stage of illness. Rather, we engaged users, including those with cognitive limitations as evidenced by UBACC scores, and other knowledge holders who could provide feedback relevant to users with varying degrees of cognitive ability. This approach allowed us to capture multiple voices and perspectives to ensure that we are designing for all persons living with dementia, including those often not considered during the design stage due to significant cognitive impairment.

Personalization and Variety of Experiences

Consistent with previous research on technology-based interventions with persons living with dementia [56,57] and activity engagement in them [58], personalization and variety of experiences emerged as key design features for our intervention. This feedback is not surprising, given the heterogeneity of identities, preferences, and lived experiences of dyads managing dementia. Increasing the number of novel or rotating visuals and sounds in each module allowed for personalization and helped to reduce boredom with repeated use. Our design team was cautious to only make slight variations in activities (eg, sounds or storytelling prompts) to reduce boredom without increasing confusion. Consistently, participants noted that nature-based sounds and animation, such as waves or chirping birds, evoked positive emotions. Therefore, our design team focused on nature scenes for many of our modules. Previous research with persons living with dementia supports that nature sounds introduced through technology, referred to as a soundscape, are perceived positively with minimal safety concerns [59,60].

Optimizing Autonomy and Choice While Sustaining Ease of Use

Feedback related to personalization highlighted an important component of designing in dementia: optimizing autonomy and choice while sustaining ease of use. In addition to increasing acceptability, the ability to tailor interventions to personal preferences may create an opportunity for persons living with dementia to have more control over their user experience. This was important for both members of the dyad to promote a sense of accomplishment, autonomy, and confidence and, ideally, allow the dyad to break the care recipient-caregiver script, enabling both members to feel as though they have ownership of their own experiences with the intervention. Previous technology-based interventions developed for persons living with dementia with mild cognitive impairments also note the importance of prioritizing user control [48]. Furthermore, in the context of dyads, increasing autonomy and choice for persons living with dementia means reducing the need for the caregiver to "deliver" the intervention. While refinements to our prototype are still needed to fully achieve this outcome, it is an important endeavor for dyadic interventionists in the context of dementia, as allowing for a break in the care recipient-caregiver script may produce positive emotional benefits for both dyad members across all stages of illness.

Users consistently provided feedback that some user-driven choices in our platform were too open ended (eg, coming up with a poem without prompts), which could lead to confusion and produce the opposite emotional experience for dyad members than intended (ie, frustration with being "wrong" and the perceived need for caregiver-directed behaviors). To address this balance, our design team attempted to provide ample instructions and feedback for use while not imposing "rules" (ie, "error-free" design) and promoting exploration. For example, several of our modules initially present users with directions for navigating the activity and then provide participants with open-ended (ie, pictures of different scenes) and forced-choice (ie, multiple choices for storytelling) decisions to guide their experience based on their needs and preferences. In addition to providing directions throughout the module, as recommended by users, we developed responsive features to reduce cognitive load and avoid frustration. For example, following effective communication strategies in the context of dementia [61], if a dyad did not respond to a multiple-choice prompt within a period of time, a forced-choice option appeared (ie, this or that). Most users with dementia rated that they felt in control of their own experience and confident in using the Isle of TEND, which may be a product of these design features.

Maintaining Engagement Among Cognitively Mismatched Users

We set out to design an IVET intervention that can maintain sustained attention and engagement among cognitively mismatched dyad users while at the same time not overwhelming the persons living with dementia or underwhelming the caregivers. Participants recommended that to achieve this goal, we needed to have a clear structure, continuous feedback, and novel stimuli, which are consistent with theory (eg, flow) [62]. In response to these recommendations, our design team developed a home screen with an introduction to the IVET intervention, tutorials and instructions for each module, and a navigation guide that appears constantly. Furthermore, our design team developed several modules that are reactive to touch or movement. For example, 1 module, the Art Garden, accesses the device camera and uses hand movement as "brushstrokes" to paint a picture. Despite these features designed to promote sustained attention and engagement, user feedback in the workshops was still mixed regarding the ability to capture and

sustain engagement across cognitively mismatched dyads and reduce caregiver-directed behaviors (eg, caregiver reiterating directions). However, it is important to note that 82% (9/11) of caregivers and 100% (9/9) of persons living with dementia still found the intervention to be entertaining. The field still needs to address the important topic of designing for mismatched cognitive abilities, as there are potential clinical benefits associated with shared engagement for person living with dementia–caregiver dyads managing dementia throughout the course of illness (eg, emotional closeness and relaxation). It is possible that design teams need to attend to more nuanced interactions between users and platforms to evaluate common elements of sustained attention and engagement across users with varying cognitive abilities (eg, movement).

Not Exacerbating Behavioral and Psychiatric Symptoms of Dementia

Regarding safety concerns, participants flagged specific stimuli that could potentially lead to distress or exacerbate behavioral or psychiatric symptoms that accompany certain types of dementia. In particular, participants recommended avoiding abstract stimuli, bright colors, or moving objects against a dark background, as these could potentially increase adverse events (eg, agitation) in more advanced stages of dementia. These images may relate to overstimulation or confusion, which increases behavioral and psychiatric disturbances in the context of advanced dementia when persons living with dementia have increased vulnerability to environmental stressors [63]. Users suggested that, in general, simpler and easily recognizable images may be better received by both persons living with dementia and caregivers. Participants also identified dark or shaded visuals as potentially problematic (eg, blurriness and glare) for users with age-related or other visual impairments. In response to the feedback, the design team removed visuals and modules that relied on limited lighting and highly abstract images. Of note, there were no observed exacerbations of dementia-related behavioral or psychiatric symptoms during workshops.

Limitations

This initial intervention development study has several limitations to consider. First, our convenience sample of caregivers, dyads, and clinicians consisted of a majority of White participants. In future pilot testing, we will ensure that participants are more representative of the US population. Second, we did not administer cognitive assessments to describe the level of cognitive impairment of our sample. Therefore, we cannot empirically comment on the cognitive profile of participants with objective measurement, including the severity of cognitive impairment. This study used self-reports from persons living with dementia and caregivers to capture the presence of dementia, and the UBACC provides additional insight into the presence of cognitive limitations inhibiting the ability to provide autonomous consent. Indeed, dementia research must balance participant burden, the reality of difficulties obtaining a dementia diagnosis, and study goals. In line with our broader goal of offering TEND as a free intervention for individuals in community settings, we included participants with varying levels of cognitive ability, as reflected by the fact that only 17% (2/12) of our sample was able to provide independent consent. To circumvent challenges with reporting on quantitative surveys for those with more advanced dementia, our team took multiple steps to support persons living with dementia in reporting their own outcomes, including using binary scales and reading questions aloud. However, these methods may have reduced the validity of our measures, and we still cannot be sure that all items were fully comprehended. Finally, we chose RDA over other forms of qualitative analysis. For this study, prioritizing a quick summary of information and relaying this to our design team consistent with RDA was necessary, but this analytic approach may also be more susceptible to researcher bias.

Conclusions

The result of this multiphasic user-centered design study with an academic-industry design team was the development of an initial dyadic IVET intervention for persons living with dementia and their caregivers to use together at home. On the basis of our conceptual model, we aimed to promote sustained attention, positive emotions, and active engagement in both dyad members to increase relationship satisfaction and reduce psychosocial distress. In this study, we highlight early intervention development with a focus on designing for usability and safety. While some of our design considerations may be helpful for future investigators and design teams focusing on dyads managing dementia, we also recognize that there are gaps that still need to be filled, such as promoting mutual engagement among cognitively mismatched dyads.

Conflicts of Interest

MP is the founder and chief executive officer of Studio Elsewhere and receives financial compensation for her role. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Example questions from phases 1 and 2 and rapid data analysis template.
[DOCX File , 27 KB - aging_v8i1e66212_app1.docx]

Multimedia Appendix 2
Toolkit for Experiential Well-Being in Dementia (Isle of TEND) modules.
[DOCX File , 12522 KB - aging_v8i1e66212_app2.docx]

References

1. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
2. Leung DK, Chan WC, Spector A, Wong GH. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2021 Sep 04;36(9):1330-1344. [doi: 10.1002/gps.5556] [Medline: 33905138]
3. Blandin K, Pepin R. Dementia grief: a theoretical model of a unique grief experience. Dementia (London) 2017 Jan;16(1):67-78 [FREE Full text] [doi: 10.1177/1471301215581081] [Medline: 25883036]
4. Gitlin LN, Jutkowitz E, Gaugler JE. Caregiver intervention research: now and into the future. College of Nursing and Health Professions. 2019. URL: https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_196527.pdf [accessed 2025-05-13]
5. Conway ER, Watson B, Tatangelo G, McCabe M. Is it all bleak? A systematic review of factors contributing to relationship change in dementia. Int Psychogeriatr 2018 Nov;30(11):1619-1637 [FREE Full text] [doi: 10.1017/S1041610218000303] [Medline: 29667571]
6. Plys E, Szapary CL, Vranceanu AM. Flexibly incorporating the dyad into psychosocial intervention: considerations for neuropalliative care. Semin Neurol 2024 Oct;44(5):559-566. [doi: 10.1055/s-0044-1791568] [Medline: 39362315]
7. Wilson SJ, Novak JR, Yorgason JB, Martire LM, Lyons KS. New opportunities for advancing dyadic health science in gerontology. Gerontologist 2024 Jan 01;64(1):gnac187 [FREE Full text] [doi: 10.1093/geront/gnac187] [Medline: 36534908]
8. Guterman EL, Allen IE, Josephson SA, Merrilees JJ, Dulaney S, Chiong W, et al. Association between caregiver depression and emergency department use among patients with dementia. JAMA Neurol 2019 Oct 01;76(10):1166-1173 [FREE Full text] [doi: 10.1001/jamaneurol.2019.1820] [Medline: 31282955]
9. Campbell JL, Rowe MA, Marsiske M. Behavioral symptoms of dementia: a dyadic effect of caregivers' stress process? Res Gerontol Nurs 2011 Jul;4(3):168-184 [FREE Full text] [doi: 10.3928/19404921-20100901-02] [Medline: 20873693]
10. Miller LM, Kaye JA, Lyons KS, Lee CS, Whitlatch CJ, Caserta MS. Well-being in dementia: a cross-sectional dyadic study of the impact of multiple dimensions of strain on persons living with dementia and their family care partners. Int Psychogeriatr 2019 May;31(5):617-626 [FREE Full text] [doi: 10.1017/S104161021800203X] [Medline: 30722800]
11. Chen KH, Merrilees J, Brown CL, Yee C, Sapozhnikova A, Wells JL, et al. Interpersonal linkage in positive and negative emotional behaviors, emotional well-being, and physical functioning in dementia caregivers. Clin Psychol Sci 2024 Jan 02;12(6):1039-1055 [FREE Full text] [doi: 10.1177/21677026231214530]
12. Hoel V, Koh WQ, Sezgin D. Enrichment of dementia caregiving relationships through psychosocial interventions: a scoping review. Front Med (Lausanne) 2022;9:1069846 [FREE Full text] [doi: 10.3389/fmed.2022.1069846] [Medline: 36687423]
13. Petrovsky DV, Yildiz M, Yefimova M, Sefcik JS, Baker ZG, Ma KP, et al. Shared activities as a protective factor against behavioral and psychological symptoms of dementia and caregiver stress. Innov Aging 2024;8(5):igae034 [FREE Full text] [doi: 10.1093/geroni/igae034] [Medline: 38660116]
14. Roland KP, Chappell NL. Meaningful activity for persons with dementia: family caregiver perspectives. Am J Alzheimers Dis Other Demen 2015 Sep;30(6):559-568 [FREE Full text] [doi: 10.1177/1533317515576389] [Medline: 25788432]
15. Van't Leven N, de Lange J, Prick AE, Pot AM. How do activating interventions fit the personal needs, characteristics and preferences of people with dementia living in the community and their informal caregivers? Dementia (London) 2019 Jan;18(1):157-177. [doi: 10.1177/1471301216662378] [Medline: 27509919]
16. Lai FH, Yan EW, Tsui WS, Yu KK. A randomized control trial of activity scheduling for caring for older adults with dementia and its impact on their spouse care-givers. Arch Gerontol Geriatr 2020;90:104167. [doi: 10.1016/j.archger.2020.104167] [Medline: 32663699]
17. Balvert SC, Del Sordo GC, Milders MV. The efficacy of dyadic interventions for community-dwelling people with dementia and their caregivers: a systematic review and meta-analysis. Ageing Res Rev 2024 Apr;96:102258. [doi: 10.1016/j.arr.2024.102258] [Medline: 38479479]
18. Gitlin LN, Marx K, Piersol CV, Hodgson NA, Huang J, Roth DL, et al. Effects of the tailored activity program (TAP) on dementia-related symptoms, health events and caregiver wellbeing: a randomized controlled trial. BMC Geriatr 2021 Oct 20;21(1):581 [FREE Full text] [doi: 10.1186/s12877-021-02511-4] [Medline: 34670502]
19. Gitlin LN, Winter L, Vause Earland T, Adel Herge E, Chernett NL, Piersol CV, et al. The Tailored Activity Program to reduce behavioral symptoms in individuals with dementia: feasibility, acceptability, and replication potential. Gerontologist 2009 Jun;49(3):428-439 [FREE Full text] [doi: 10.1093/geront/gnp087] [Medline: 19420314]
20. Motejlek J, Alpay E. Taxonomy of virtual and augmented reality applications in education. IEEE Trans Learn Technol 2021 Jun 1;14(3):415-429. [doi: 10.1109/tlt.2021.3092964]
21. Smith JW. Immersive virtual environment technology to supplement environmental perception, preference and behavior research: a review with applications. Int J Environ Res Public Health 2015 Sep 11;12(9):11486-11505 [FREE Full text] [doi: 10.3390/ijerph120911486] [Medline: 26378565]

22. Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL, Bailenson JN. Immersive virtual environment technology as a methodological tool for social psychology. Psychol Inquiry 2002 Aug;13(2):103-124 [FREE Full text] [doi: 10.1207/s15327965pli1302_03]
23. Cheng C, Baker G, Dursun S. Use of multisensory stimulation interventions in the treatment of major neurocognitive disorders. Psychiatry Clin Psychopharmacol 2019 Dec 23;29(4):916-921. [doi: 10.1080/24750573.2019.1699738]
24. Burdea GC, Coiffet P. Virtual Reality Technology. Hoboken, NJ: John Wiley & Sons; 2003.
25. Li H, Ding Y, Zhao B, Xu Y, Wei W. Effects of immersion in a simulated natural environment on stress reduction and emotional arousal: a systematic review and meta-analysis. Front Psychol 2022 Jan 9;13:1058177 [FREE Full text] [doi: 10.3389/fpsyg.2022.1058177] [Medline: 36698558]
26. Shu S, Woo BK. Use of technology and social media in dementia care: current and future directions. World J Psychiatry 2021 Apr 19;11(4):109-123 [FREE Full text] [doi: 10.5498/wjp.v11.i4.109] [Medline: 33889536]
27. Strong J. Immersive virtual reality and persons with dementia: a literature review. J Gerontol Soc Work 2020 Apr;63(3):209-226. [doi: 10.1080/01634372.2020.1733726] [Medline: 32091323]
28. Hirt J, Beer T. Use and impact of virtual reality simulation in dementia care education: a scoping review. Nurse Educ Today 2020 Jan;84:104207. [doi: 10.1016/j.nedt.2019.104207] [Medline: 31669968]
29. Hodge J, Balaam M, Hastings S, Morrissey K. Exploring the design of tailored virtual reality experiences for people with dementia. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC. [doi: 10.1145/3173574.3174088]
30. Kokorelias KM, Chiu M, Paul S, Zhu L, Choudhury N, Craven CG, et al. Use of virtual reality and augmented reality technologies to support resilience and skill-building in caregivers of persons with dementia: a scoping review. Cureus 2024 Jul;16(7):e64082. [doi: 10.7759/cureus.64082] [Medline: 39114214]
31. Huang Y, Ho KH, Christensen M, Wong DW, Wang S, Su JJ, et al. Virtual reality-based simulation intervention for enhancing the empathy of informal caregivers of people with dementia: a mixed-methods systematic review. Int J Ment Health Nurs 2024 Apr;33(2):241-258. [doi: 10.1111/inm.13240] [Medline: 37817470]
32. Zhi S, Gao S, Sun J, Ma D, Zhao Y, He M, et al. Virtual reality-based dementia educational programmes for formal and informal caregivers of people with dementia: a scoping review. J Clin Nurs 2024 Apr 29. [doi: 10.1111/jocn.17198] [Medline: 38685766]
33. Goodall G, Ciobanu I, Taraldsen K, Sørgaard J, Marin A, Draghici R, et al. The use of virtual and immersive technology in creating personalized multisensory spaces for people living with dementia (SENSE-GARDEN): protocol for a multisite before-after trial. JMIR Res Protoc 2019 Aug 19;8(9):e14096 [FREE Full text] [doi: 10.2196/14096] [Medline: 31538942]
34. Sideman AB, Merrilees J, Dulaney S, Kiekhofer R, Braley T, Lee K, et al. "Out of the clear blue sky she tells me she loves me": connection experiences between caregivers and people with dementia. J Am Geriatr Soc 2023 Jul;71(7):2172-2183 [FREE Full text] [doi: 10.1111/jgs.18297] [Medline: 36917135]
35. Paliokas I, Tzallas A, Katertsidis N, Votis K, Tzovaras D. Gamification in social networking: a platform for people living with dementia and their caregivers. In: Proceedings of the 17th International Conference on Bioinformatics and Bioengineering. 2017 Presented at: BIBE 2017; October 23-25, 2017; Washington, DC. [doi: 10.1109/bibe.2017.00015]
36. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. Clin Psychol (New York) 2016 Jun;23(2):180-200 [FREE Full text] [doi: 10.1111/cpsp.12154] [Medline: 29456295]
37. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Commun ACM 1985 Mar;28(3):300-311. [doi: 10.1145/3166.3170]
38. Witteman HO, Dansokho SC, Colquhoun H, Coulter A, Dugas M, Fagerlin A, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev 2015 Jan 26;4(1):11 [FREE Full text] [doi: 10.1186/2046-4053-4-11] [Medline: 25623074]
39. Ceccacci S, Generosi A, Giraldi L, Mengoni M. An user-centered approach to design smart systems for people with dementia. In: Proceedings of the 2017 IEEE 7th International Conference on Consumer Electronics - Berlin. 2017 Presented at: ICCE-Berlin 2017; September 3-6, 2017; Berlin, Germany. [doi: 10.1109/icce-berlin.2017.8210650]
40. National Institutes of Health (NIH) homepage. National Institutes of Health (NIH). URL: https://www.nih.gov/ [accessed 2024-08-07]
41. Onken L. Implementation science at the national institute on aging: the principles of it. Public Policy Aging Report 2022;32(1):39-41 [FREE Full text] [doi: 10.1093/ppar/prab034]
42. Rochon EA, Sy M, Phillips M, Anderson E, Plys E, Ritchie C, et al. Bio-experiential technology to support persons with dementia and care partners at home (TEND): protocol for an intervention development study. JMIR Res Protoc 2023 Dec 29;12:e52799 [FREE Full text] [doi: 10.2196/52799] [Medline: 38157239]
43. Duncan E, O'Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]
44. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]

45. Seaman JB, Terhorst L, Gentry A, Hunsaker A, Parker LS, Lingler JH. Psychometric properties of a decisional capacity screening tool for individuals contemplating participation in Alzheimer's disease research. J Alzheimers Dis 2015;46(1):1-9 [FREE Full text] [doi: 10.3233/JAD-142559] [Medline: 25765917]
46. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996.
47. Miguel Cruz A, Daum C, Comeau A, Salamanca JD, McLennan L, Neubauer N, et al. Acceptance, adoption, and usability of information and communication technologies for people living with dementia and their care partners: a systematic review. Disabil Rehabil Assist Technol 2023 May;18(4):443-457. [doi: 10.1080/17483107.2020.1864671] [Medline: 33378627]
48. Quintana M, Anderberg P, Sanmartin Berglund J, Frögren J, Cano N, Cellek S, et al. Feasibility-usability study of a tablet app adapted specifically for persons with cognitive impairment-SMART4MD (support monitoring and reminder technology for mild dementia). Int J Environ Res Public Health 2020 Sep 18;17(18):6816 [FREE Full text] [doi: 10.3390/ijerph17186816] [Medline: 32961993]
49. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
50. Kolanowski A, Gilmore-Bykovskyi A, Hill N, Massimo L, Mogle J. Measurement challenges in research with individuals with cognitive impairment. Res Gerontol Nurs 2019 Jan 01;12(1):7-15 [FREE Full text] [doi: 10.3928/19404921-20181212-06] [Medline: 30653646]
51. Schexnayder J, Perry KR, Sheahan K, Majette Elliott N, Subramaniam S, Strawbridge E, et al. Team-based qualitative rapid analysis: approach and considerations for conducting developmental formative evaluation for intervention design. Qual Health Res 2023 Jul 06;33(8-9):778-789. [doi: 10.1177/10497323231167348] [Medline: 37278662]
52. Halcomb EJ, Davidson PM. Is verbatim transcription of interview data always necessary? Appl Nurs Res 2006 Feb;19(1):38-42. [doi: 10.1016/j.apnr.2005.06.001] [Medline: 16455440]
53. Suchman L, Omoluabi E, Kramer J, Vallin J, Sedlander E, Gitome S, et al. Analyzing fast and slow: combining traditional and rapid qualitative analysis to meet multiple objectives of a complex transnational study. Front Sociol 2023;8:961202 [FREE Full text] [doi: 10.3389/fsoc.2023.961202] [Medline: 36818663]
54. Vindrola-Padros C, Johnson GA. Rapid techniques in qualitative research: a critical review of the literature. Qual Health Res 2020 Aug;30(10):1596-1604. [doi: 10.1177/1049732320921835] [Medline: 32667277]
55. Lewinski AA, Crowley MJ, Miller C, Bosworth HB, Jackson GL, Steinhauser K, et al. Applied rapid qualitative analysis to develop a contextually appropriate intervention and increase the likelihood of uptake. Med Care 2021 Jun 01;59(Suppl 3):S242-S251 [FREE Full text] [doi: 10.1097/MLR.0000000000001553] [Medline: 33976073]
56. Goodall G, Taraldsen K, Serrano JA. The use of technology in creating individualized, meaningful activities for people living with dementia: a systematic review. Dementia (London) 2021 May;20(4):1442-1469 [FREE Full text] [doi: 10.1177/1471301220928168] [Medline: 32475166]
57. Eun SJ, Kim EJ, Kim J. Artificial intelligence-based personalized serious game for enhancing the physical and cognitive abilities of the elderly. Future Gener Comput Syst 2023 Apr;141:713-722. [doi: 10.1016/j.future.2022.12.017]
58. Trahan MA, Kuo J, Carlson MC, Gitlin LN. A systematic review of strategies to foster activity engagement in persons with dementia. Health Educ Behav 2014 Oct;41(1 Suppl):70S-83S [FREE Full text] [doi: 10.1177/1090198114531782] [Medline: 25274714]
59. Houben M, Brankaert R, Bakker S, Kenning G, Bongers I, Eggen B. The role of everyday sounds in advanced dementia care. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu, HI. [doi: 10.1145/3313831.3376577]
60. Devos P, Aletta F, Thomas P, Petrovic M, Vander Mynsbrugge T, Van de Velde D, et al. Designing supportive soundscapes for nursing home residents with dementia. Int J Environ Res Public Health 2019 Dec 04;16(24):4904 [FREE Full text] [doi: 10.3390/ijerph16244904] [Medline: 31817300]
61. Small JA, Gutman G, Makela S, Hillhouse B. Effectiveness of communication strategies used by caregivers of persons with Alzheimer's disease during activities of daily living. J Speech Lang Hear Res 2003 Apr;46(2):353-367. [doi: 10.1044/1092-4388(2003/028)] [Medline: 14700377]
62. Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi. Dordrecht, The Netherlands: Springer; 2014.
63. Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. BMJ 2015 Mar 02;350:h369 [FREE Full text] [doi: 10.1136/bmj.h369] [Medline: 25731881]

Abbreviations

Isle of TEND: Toolkit for Experiential Well-Being in Dementia
IVET: immersive virtual environment technology
NIH: National Institutes of Health
RDA: rapid data analysis
SUS: System Usability Scale
UBACC: University of California, San Diego Brief Assessment of Capacity to Consent Questionnaire

Edited by M Antoniou; submitted 24.09.24; peer-reviewed by E Tonkin, D Urbanski, R Bratches; comments to author 06.01.25; revised version received 17.03.25; accepted 30.03.25; published 21.05.25.

Please cite as:
Rochon EA, Thacker A, Phillips M, Ritchie C, Vranceanu AM, Plys E
Developing a Dyadic Immersive Virtual Environment Technology Intervention for Persons Living With Dementia and Their Caregivers: Multiphasic User-Centered Design Study
JMIR Aging 2025;8:e66212
URL: https://aging.jmir.org/2025/1/e66212
doi:10.2196/66212
PMID:

Original Paper

Design of a Mobile App and a Clinical Trial Management System for Cognitive Health and Dementia Risk Reduction: User-Centered Design Approach

Hannes Hilberger[1,2], MSc[‡]; Bianca Buchgraber-Schnalzer[1], MSc; Simone Huber[1], MSc; Theresa Weitlaner[1], MSc; Markus Bödenler[1], PhD; Alara Abaci[3], MSc; Jeroen Bruinsma[4], PhD; Ana Diaz[5], PhD; Anna Giulia Guazzarini[6,7], MSc; Jenni Lehtisalo[8], PhD; Seungjune Lee[9,10]; Vasileios Loukas[11,12], MSc; Francesca Mangialasche[3,13], PhD; Patrizia Mecocci[6,7], PhD; Tiia Ngandu[3,8,14], PhD; Anna Rosenberg[3,8], PhD; Elisabeth Stögmann[9,10], MD; Konsta Valkonen[8,15], MSc; Elena Uhlik[9,10]; Helena Untersteiner[9,10], MD; Laura Kneß[1], MSc; Helmut Ahammer[2], PhD; Sten Hanke[1], PhD

[1]Institute of eHealth, University of Applied Sciences - FH JOANNEUM, Graz, Austria

[2]Gottfried Schatz Research Center, Division of Medical Physics and Biophysics, Medical University of Graz, Graz, Austria

[3]Division of Clinical Geriatrics, Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[4]Department of Health Promotion, Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands

[5]Alzheimer Europe, Luxembourg, Luxembourg

[6]Division of Gerontology and Geriatrics, Department of Medicine and Surgery, University of Perugia, Perugia, Italy

[7]Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[8]Lifestyles and Living Environments Unit, Department of Public Health, Finnish Institute for Health and Welfare, Helsinki, Finland

[9]Department of Neurology, Medical University of Vienna, Vienna, Austria

[10]Comprehensive Center for Clinical Neurosciences and Mental Health, Medical University of Vienna, Vienna, Austria

[11]Department of Materials Science and Engineering, Unit of Medical Technology and Intelligent Information Systems, University of Ioannina, Ioannina, Greece

[12]Biomedical Research Institute, Foundation for Research and Technology–Hellas, Ioannina, Greece

[13]Medical Unit Aging, Theme Inflammation and Aging, Karolinska Institutet, Stockholm, Sweden

[14]Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

[15]Institute of Clinical Medicine, University of Eastern Finland, Kuopio, Finland

[‡]LETHE Consortium

Corresponding Author:
Hannes Hilberger, MSc
Institute of eHealth
University of Applied Sciences - FH JOANNEUM
Institute of eHealth
Eckertstraße 30i
Graz, 8020
Austria
Phone: 43 316 5453 6529
Email: hannes.hilberger@fh-joanneum.at

Abstract

Background: The rising prevalence of dementia is a major concern, with approximately 45% of cases linked to 14 modifiable risk factors. The European project LETHE aims to develop a personalized digital intervention model to delay or prevent cognitive decline through risk factor management.

Objective: The objective of our study was to design a clinical trial platform for older individuals at risk of cognitive decline, including a mobile app for study participants and a clinical trial management system (CTMS) for health professionals.

Methods: Using a user-centered design approach, workshops and feedback rounds involved potential participants representing the target group and professionals. The LETHE app's usability was assessed among 156 older adults enrolled in a 2-year

multinational randomized controlled trial evaluating the feasibility of a digitally supported lifestyle program for dementia risk reduction. The randomized controlled trial is currently ongoing; the System Usability Scale (SUS) was administered 1 month after baseline to map first user experiences. Feedback on the LETHE CTMS was collected from 21 users.

Results: Of the 78 participants in the trial intervention group, 66 (85%) provided responses for the mobile app, with a median SUS score of 70 (IQR 55-82). Within the control group, 73% (57/78) of responses were received, with a median SUS score of 73 (IQR 63-90). For the CTMS, we received 71% (15/21) of responses, and the feedback was mostly positive. A ranking of the features that could be considered beyond state of the art showed that the integration of personalized activities (mean 2.23, SD 1.17) and real-time appointments (mean 2.46, SD 1.51) were considered the most novel ones.

Conclusions: The LETHE app and CTMS were developed to support a personalized digital intervention method within a study involving 156 participants. Limitations include participants having digital literacy and internet access, potentially impacting the generalizability of the findings. Despite these limitations, positive feedback and high usability scores suggest promising potential for the LETHE app and CTMS in supporting personalized interventions to prevent cognitive decline in older adults.

(JMIR Aging 2025;8:e66660) doi:10.2196/66660

KEYWORDS

clinical trial; medical informatics; health information systems; eHealth; mobile health; mHealth; mobile apps; dementia; cognitive decline; prevention; multidomain interventions; artificial intelligence; AI

Introduction

Background

Overview

The rising prevalence of dementia, driven by sociodemographic changes, constitutes a significant global health challenge. Projections indicate a notable rise from 57 million affected individuals in 2019 to an anticipated 153 million by 2050 [1]. Notably, the Lancet Commission's report underscores that approximately 45% of all cases of dementia are associated with 14 potentially modifiable risk factors at different phases of the life span [2], and diet has also been suggested as an additional factor [3]. Previous multidomain intervention studies have demonstrated their positive impact on cognitive performance while simultaneously targeting various lifestyle domains, including diet, physical activity, cognitive training, management of cardiovascular risk factors, or social interaction [4-6].

Information and communications technology (ICT) solutions such as apps running on phones and tablets, as well as wearable devices that collect user-generated health data (digital biomarkers) in an automated way, could potentially help provide digitally supported interventions for a broader audience and groups at risk of cognitive decline. At the same time, these technologies can collect data to monitor progress and adherence to lifestyle interventions. Other research projects have shown that digital interventions can be effective and feasible for older adults, but this depends on several factors such as digital literacy, usability of the technology, and the design of the study with regard to human support (hybrid intervention) [7-9].

The Horizon 2020 LETHE project [10] has been initiated to provide such a personalized digital intervention model and components for reducing risk of developing cognitive decline by evolving the successful Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) multidomain intervention model [5]. Leveraging ICT-based methods and guided in-person as well as remote sessions, LETHE aims to prevent cognitive decline in at-risk older individuals by providing personalized feedback and intervention methods.

The LETHE project focuses on the aforementioned modifiable risk factors grouped into the following lifestyle domains: "Physical Activity," "Nutrition," "Cognitive Training," "Management of vascular/metabolic risk factors," "Sleep/Relaxation," and "Social Activity."

The LETHE study encompasses a currently ongoing 2-year randomized controlled multicenter, parallel-group feasibility trial (ClinicalTrials.gov; NCT05565170) involving 156 participants at risk of cognitive decline distributed across clinical centers in Austria (Medical University of Vienna), Finland (Finnish Institute for Health and Welfare), Sweden (Karolinska Institute), and Italy (University of Perugia). Participants were equally randomized to one of the study groups (intervention and control), both supported by digital tools (mobile app and smartwatch). At baseline, participants had a mean age of 68.8 (SD 4.5) years, with an average of 14.9 (SD 3.1) years of education. A total of 64.7% of the participants were women. Notably, most participants (84%) used their smartphones at least 6 times per day. Two-thirds of the participants had previous experience using their smartphones for health tracking, and approximately 48% had used lifestyle-related apps. In addition, 63% of the participants reported using the internet for eHealth-related tasks such as booking physician appointments or checking test results [11].

To provide intervention components and continuously collect data from the trial participants, a LETHE smartphone app was designed and implemented. Clinical data, including blood test results, are entered into the LETHE clinical trial management system (CTMS), which was also developed during the course of the project, by health professionals at each visit. The CTMS was developed to ensure seamless integration with the app, allowing all participant-entered data to be displayed for health professionals to review and respond accordingly. Furthermore, data and services from third parties have been integrated, including cognitive training data and data from Fitbit devices, which track steps, sleep, and physical activity. The feasibility

of other novel technologies (audio glasses and a robot that reflect the content of the smartphone app) is being investigated in a substudy.

This study aimed to describe the design and implementation of the LETHE technical components, which were carefully developed through a joint process involving potential users, clinical experts, and designers and technicians.

Hybrid Intervention Approach

LETHE adopts a hybrid intervention (blended therapies) approach, seamlessly integrating digital intervention sessions with in-person and group sessions led by a trained coach (eg, physiotherapist or nutritionist). This comprehensive strategy, meticulously managed by health professionals, is in harmony with the study's mobile app and CTMS concept.

The literature on hybrid interventions for older adults at risk of cognitive decline, incorporating technology with coaching and peer sessions in comparison to stand-alone technology interventions, is still developing and has so far mainly focused on already symptomatic individuals. Nevertheless, emerging studies indicate that fostering social connection and interaction can be pivotal for success in dementia interventions [12]. Digital health platforms, easily accessible via mobile technology, can be efficient and cost-effective tools for dementia prevention. They not only offer individualized cognitive training but can also provide an engaging user experience [13,14]. The currently ongoing prevention of dementia using mobile phone applications (PRODEMOS) study creates an evidence-based dementia prevention strategy using mobile health (mHealth) accessible to at-risk individuals [15]. In contrast, Maintain Your Brain was a completely online lifestyle intervention targeting modifiable risk factors for dementia that was delivered to the participants via a web interface [16]. Wesselman et al [17] conducted a comprehensive overview of web-based multidomain lifestyle programs, combining it with a meta-analysis to evaluate their effectiveness. Through a systematic literature research, they collected a wide range of web-based programs.

Other studies have provided app-supported self-regulation for older adults based on self-determination theory, with a tablet facilitating tailored exercise programs and playing a key role in action planning and execution of behaviors [18]. The involvement of a personal coach was pivotal, adapting exercises to individual preferences and providing motivating remote monitoring to empower older adults to enhance physical activity levels at home. The experience conveyed in the study by Mehra et al [18] emphasizes the value of the personal coach, particularly during the initial phases of goal setting, behavior execution, and evaluation in self-regulation to achieve specific goals by guiding one's own behavior. Their findings indicate that the availability of a personal coach remains crucial even when technology supports the intervention. Nevertheless, concerns about adherence and safety arose in the absence of instructor guidance [18].

The success of this hybrid intervention approach could be further supported by the seamless interaction of a mobile app and a web-based CTMS for professionals. This integrated system allows professionals to interact with and monitor patients' development, enabling timely interventions when needed.

Digital Intervention Applications

mHealth apps are increasingly used by individuals to engage in health behaviors, aid in the self-management of chronic conditions, or enact preventive measures [19]. These apps hold promise in personalizing and tailoring behavior change interventions based on real-time data, thereby enhancing health outcomes [20]. Another crucial consideration is the necessity for co-designing apps in collaboration with the end users and health professionals from the relevant sector [21]. This is particularly vital for older adults as their participation in the technological design process is pivotal for their acceptance and adoption of the technology [22].

Digital interventions, whether delivered through mHealth apps or computers, are subject to evaluation in randomized controlled trials (RCTs) targeting various conditions such as psychotic disorders [23,24], mental health [25-28], eating disorders [29], diabetes [30], chronic pain [31], insomnia [32], or speech disorders [33]. These interventions often incorporate features such as audio or video content for physical activity or relaxation [23,24,26,32]; mood tracking [23,25,26]; personal tasks [23,25,31]; questionnaires [25,27,29,31]; educational materials [23-26,32]; self-tracking of, for example, vital signs in the form of a diary [24,32]; and habit libraries for behavior changes [34].

An umbrella review comprising 48 systematic reviews concentrating on mHealth apps in RCTs across various health conditions such as diabetes or hypertension indicates the potential effectiveness of app-based health interventions [35]. This further strengthens the approach of delivering a mobile-based app for study participants.

CTMSs for Intervention Projects

Equally important when intervention apps are used in clinical trials or research settings is central monitoring and management by the study coordinator or a corresponding physician. Using a CTMS and the integration of electronic data capture has become commonplace to serve this purpose [36]. Such a platform should facilitate and effectively support the multidimensional data management process in clinical trials [37] throughout their phases, from participant onboarding to completion [38].

Electronic systems allow for streamlined data transfer from clinics to the CTMS, remote enrollment capabilities, greater transparency of trial conduct, timeline monitoring such as specific tasks, tracking of participant visits, enhanced research documentation, and robust reporting [36,38,39]. A CTMS presents a variety of advantages; for example, it enables research teams to access up-to-date study information and simplifies collaboration as all project members can work efficiently together on the same task [38]. Moving beyond traditional methods of data collection in clinical trials, which may involve manual completion of paper case report forms (CRFs), leads to increased accuracy in findings, enhanced productivity as all necessary elements for managing the trial are consolidated in a single location, and higher data quality and compliance while reducing the risk of bias in clinical outcomes [36]. Important

considerations when selecting a CTMS include the feature set, usability, customization, and cost [38].

A systematic review of 19 research papers examining the technical features of clinical data management systems revealed that most of these systems were developed on a web-based platform to meet the individual needs of specific clinical trials within a short time frame [40]. Reportedly, such systems used in research centers showed limitations and inability to fully support the automation of all dimensions of the clinical data and workflow management process. In addition, the review found that most of the systems lacked flexibility and extensibility for further system development.

Health professionals such as study nurses, who are key members of the clinical research team and one of the targeted user groups for CTMSs, play a critical role in achieving accurate outcomes for clinical research studies. Furthermore, the effective implementation of a CTMS in RCT studies can have multiple benefits, including improved completion rates and increased fidelity while ensuring the safety of individual research participants [36].

The studies mentioned in this section collectively underscore the value of CTMSs in enhancing the efficiency and effectiveness of complex multisite intervention projects.

Objectives

This paper describes the iterative design process and reports on the technical implementation of the LETHE app and the LETHE CTMS within the context of a multifaceted and hybrid intervention to prevent cognitive decline in older adults. By sharing the results, we aim to contribute valuable insights to the field of dementia prevention and risk factor management through personalized ICT-supported hybrid intervention methods. We collaborated intensively with end users and health professionals to design the clinical trial platform. In addition, in this study, we evaluated how the trial participants and clinicians accepted the developed applications by using subjective usability assessments. This helped us to understand whether such digital tools keep the participants more engaged and identify directions for further improvement.

Methods

Design of a Mobile App and Clinical Trial Platform

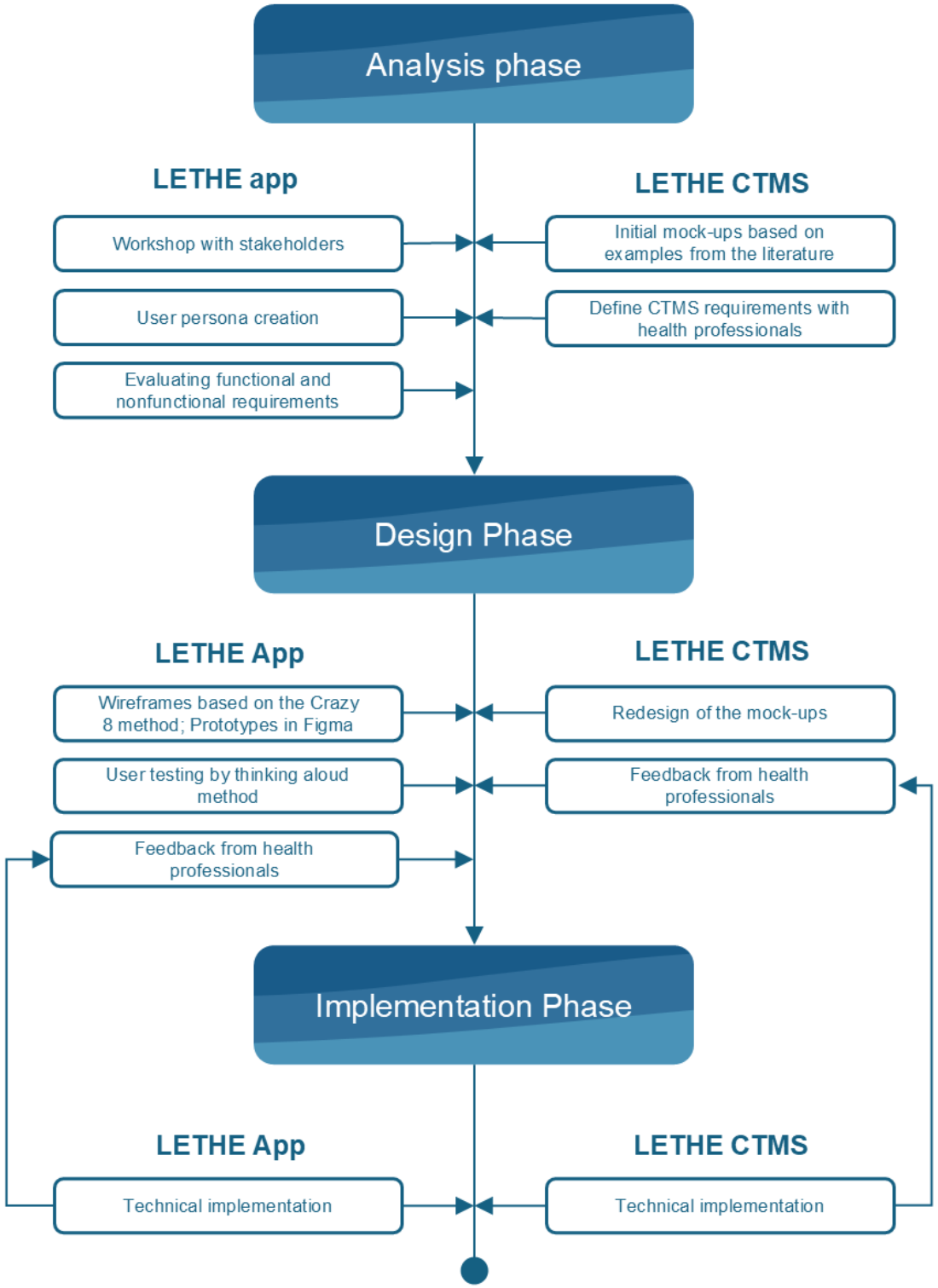
Overview

To ensure high usability, the user-centered design (UCD) approach was chosen for the LETHE app as well as for the CTMS design process. The UCD process involves the participation of end users (in our case, older individuals at risk of cognitive decline and health professionals supporting them during a preventive lifestyle intervention using the CTMS) to define key user requirements. The design approach consists of 3 main phases: analysis, design, and implementation [41].

The UCD approach encompasses 5 key principles: defining user requirements, gathering and considering feedback from users to specify requirements, involving users from the outset to evaluate design iterations, consistently adhering to UCD, and using an iterative design process [42].

Figure 1 illustrates the different design phases of the LETHE app and the LETHE CTMS.

Figure 1. Different phases for the design and development of the LETHE app and LETHE clinical trial management system (CTMS).



Requirement Analysis and Design of the LETHE App

First, a requirement workshop, which was held online, was carried out with the goal of listing requirements for the entire LETHE ecosystem and especially for the LETHE app. A total of 21 individuals from the LETHE project, including clinicians, behavioral experts, and representatives of public involvement in dementia, were invited to the workshop.

During the workshop, 3 user personas [43] were created, which represented typical end users of the LETHE app. General requirements were evaluated via open-ended questions from a clinical view. Specific requirements such as the design or the scientific use were ranked according to importance. The functional requirements of the LETHE app included, on the one hand, the content features according to the FINGER lifestyle domains; on the other hand, general digital requirements were collected from professionals during the workshop. The design

of the digital intervention was based on established behavior change theories to ensure that the app's features were effective. We focused on understanding relevant psychological factors that influence behaviors related to dementia risk [44,45] by relying on multiple behavior change theories to achieve a comprehensive understanding [46]. We used intervention mapping elements [47] and relied on a behavior change taxonomy [48] to select strategies that would best target relevant psychological factors to promote behavior change. Methods such as self-monitoring, setting clear goals, and planning for challenges were built into the app's features. Throughout the design process, a collaboration with health professionals and an advisory board of older adults at risk of or living with dementia was established. The results of this workshop were collected via online collaborating tools such as Mentimeter [49] and Padlet [50].

Afterward, wireframes based on the requirements and identified personas were created. During the user experience design, the method "The crazy 8" for wireframing was used to create as many screens as possible [51]. This method is a fast-sketching method to engage designers in sketching 8 distinct ideas in 8 minutes. In the next step, the wireframes were built using Figma (Figma, Inc) [52] as clickable prototypes to obtain continuous feedback from the involved health professionals.

In addition, a first user test with 4 German-speaking citizens aged between 65 and 85 years was carried out at the Department of Neurology at the Medical University of Vienna, and results were integrated during the implementation phase. One participant had a diagnosis of dementia, whereas the other 3 had no diagnosis. The user tests were conducted using the wireframes as clickable prototypes. First, participants explored the prototype using the thinking-aloud method [53]. Afterward, they were asked to perform 2 specific tasks (ie, choosing among a prepopulated list of pieces of advice for lifestyle improvement and setting them as goals for personal change and entering self-measured blood pressure values). Finally, participants provided verbal feedback and shared ideas and recommendations to further improve the prototypes.

When an implementation-ready design was obtained after the workshop and the user test, the implementation phase (including translation tasks) involved continuous feedback from professionals. Textbox 1 summarizes the phases for creating the design of the LETHE app.

Textbox 1. Phases for creating the design of the LETHE app, with steps 4 and 5 being recurring.

- Workshop for creation of user personas and functional and nonfunctional requirements (step 1)
- Creation of wireframes and clickable prototypes based on the workshop (step 2)
- User testing with potential end users (step 3)
- Adaption of wireframes and implementation phase (step 4)
- Continuous improvement, modification, and development after feedback from professionals and suggestions by study participants and advisory board members (step 5)

Requirement Analysis and Design of the LETHE CTMS

In the initial phase of our UCD methodology, an interactive workshop dedicated to the analysis of user requirements for the LETHE CTMS was conducted. The workshop was held online via Microsoft Teams (Microsoft Corp) with 13 health professionals as participants from different study centers located in Finland, Italy, Sweden, and Austria, as well as technical partners involved in the LETHE project. The professionals of the study centers included neurologists, gerontologists, public health experts, and professionals in the field of quality of life and dementia who represented the end users of the CTMS. The aim was to collaboratively identify the essential features and functionalities required for the LETHE CTMS. The workshop used the Mural tool [54], facilitating the creation of digital whiteboards using sticky notes for brainstorming in smaller breakout rooms consisting of 3 to 4 persons. In those breakout rooms, initial design considerations were discussed. Those considerations included general thoughts on a CTMS, an overview of the participants, data entering, and an artificial intelligence (AI) risk simulation. The process behind the design of the CTMS is described in Textbox 2 [55-57].

The results of the workshop influenced the subsequent design phase, during which multiple design proposals were iteratively developed incorporating feedback from clinical professionals. Once the design phase concluded, we transitioned into the implementation phase.

Textbox 2. Process to design the LETHE clinical trial management system, with steps 3 and 4 being recurring steps.

- Creation of initial mock-ups based on examples from the literature [55-57] and the study protocol of the LETHE trial for workshops together with clinical and technical partners (step 1)
- Conduction of the workshops to gather feedback and potentially missing features (step 2)
- Redesign of the mock-ups and technical implementation (step 3)
- Feedback on the redesign (step 4)

Evaluation of the Mobile App and Clinical Trial Platform

Overview

This section provides a detailed overview of the methodology used to analyze user perspectives within the currently ongoing 2-year clinical feasibility trial, focusing on both the LETHE app and the LETHE CTMS (or LETHE dashboard). The aim was to gather insights on basic usability assessments, encompassing considerations such as satisfaction with the applications.

The analysis was conducted using R (version 4.3.2; R Foundation for Statistical Computing) [58]. The normality test for the data was carried out using the Shapiro-Wilk test (α=.05). To assess the difference in the System Usability Scale (SUS) scores [59] between the study groups, the Mann-Whitney *U* test was used. For pairwise comparisons of the various countries, the Kruskal-Wallis test followed by the Dunn post hoc test with Bonferroni correction was used.

Evaluation of the LETHE App

Older adults who participated in the LETHE RCT provided feedback within the LETHE app through a structured questionnaire, "Experiences with LETHE app," consisting of 10 closed-ended questions based on an adapted version of the SUS, which was simplified for older adults and adults with cognitive impairment [60]. Using a Likert Scale [61] with 5 points ranging from "Strongly Disagree" to "Strongly Agree," participants expressed subjective feelings about the frequency of LETHE app use, perceptions of app complexity, and their need for assistance or confidence in navigating the mobile app. The questionnaire can be found in Multimedia Appendix 1.

This survey will be administered at 3 distinct time points throughout the study: 1 month, 6 months, and 24 months after randomization. Participants have 4 weeks to complete the survey at each time point. This timeline allows for the observation of participant sentiments over an extended duration, offering insights into the subjective aspects of use patterns and the perceived complexity of a mobile app specifically designed for older individuals. As the trial is still ongoing, this study focused on data from the first time point. As a result, we share participants' initial user experiences at this stage.

The LETHE app exists in 2 versions: one for the intervention group and one for the control group. Due to the limited functionality of the version for the control group, which lacks more sophisticated features such as personalized or intervention activities, the analysis was conducted separately for the intervention and control groups.

Within the LETHE trial, 156 study participants were invited to complete the questionnaire, with 78 (50%) in the intervention group and 78 (50%) in the control group. The calculated SUS was assigned a grade [62] as shown in Table 1, with "C" as the average grade and 68 as the center of the range.

Table 1. Interpretation of the System Usability Scale (SUS) score based on a grading scale by Sauro and Lewis [62].

SUS range	Grade	Percentile range
84.1-100	A+	96-100
80.8-84	A	90-95
78.9-80.7	A–	85-89
77.2-78.8	B+	80-84
74.1-77.1	B	70-79
72.6-74	B–	65-69
71.1-72.5	C+	60-64
65-71	C	41-59
62.7-64.9	C–	35-40
51.7-62.6	D	15-34

Evaluation of the LETHE CTMS

To obtain a comprehensive understanding of perspectives from relevant health professionals who used the LETHE CTMS, a 41-question online survey was conducted including both open- and closed-ended questions via Microsoft Forms (Microsoft Corp). The survey was distributed 10 months after the start of the study, and users were given 3 weeks to complete it. The primary objective was to gather feedback on the functionalities of the LETHE CTMS and assess user satisfaction with its implementation. The survey aimed to provide insights essential for mitigating common bottlenecks in the future development of similar tools. Notably, the survey was self-administered and structured into different parts, starting with a brief introduction outlining its objective before participants provided information about their role in the project. Subsequently, general inquiries about the CTMS, such as the onboarding process and overall experience, were posed. Participants then provided feedback on each functionality. Finally, recommendations, challenges, and additional comments were solicited.

All users of the LETHE CTMS were invited to participate and were informed about the survey duration.

The mixing of open- and closed-ended questions served a dual purpose. Closed-ended questions, including Likert scaling and ranking formats, enabled the quantification of participant sentiments, whereas parallel open-ended inquiries facilitated the collection of qualitative feedback. Noteworthy is the

inclusion of a single ranking question designed to identify the beyond-state-of-the-art features of the LETHE CTMS.

Moreover, a 5-point Likert scale [61] was used for closed-ended questions, encompassing the following response options: “Strongly Agree,” “Agree,” “Neither Agree nor Disagree,” “Disagree,” and “Strongly Disagree.”

The structure of the survey can be found in Multimedia Appendix 2.

Ethical Considerations

The trial has been approved by ethical committees in Austria (Ethics Committee of the Medical University of Vienna; 1392/2022), Finland (Hospital District of Helsinki and Uusimaa Ethical Committee; HUS/13675/2022), Italy (Regional Ethics Committee Umbria; 25723/22/AV), and Sweden (Swedish Ethical Review Authority; 2022-03961-01). All participants provided written informed consent before enrollment. The data used in this study for analysis were anonymized. Study participants received a wearable device, and a phone if they chose to have one. Some of those in the intervention group were additionally provided with a tablet for use during the trial.

Results

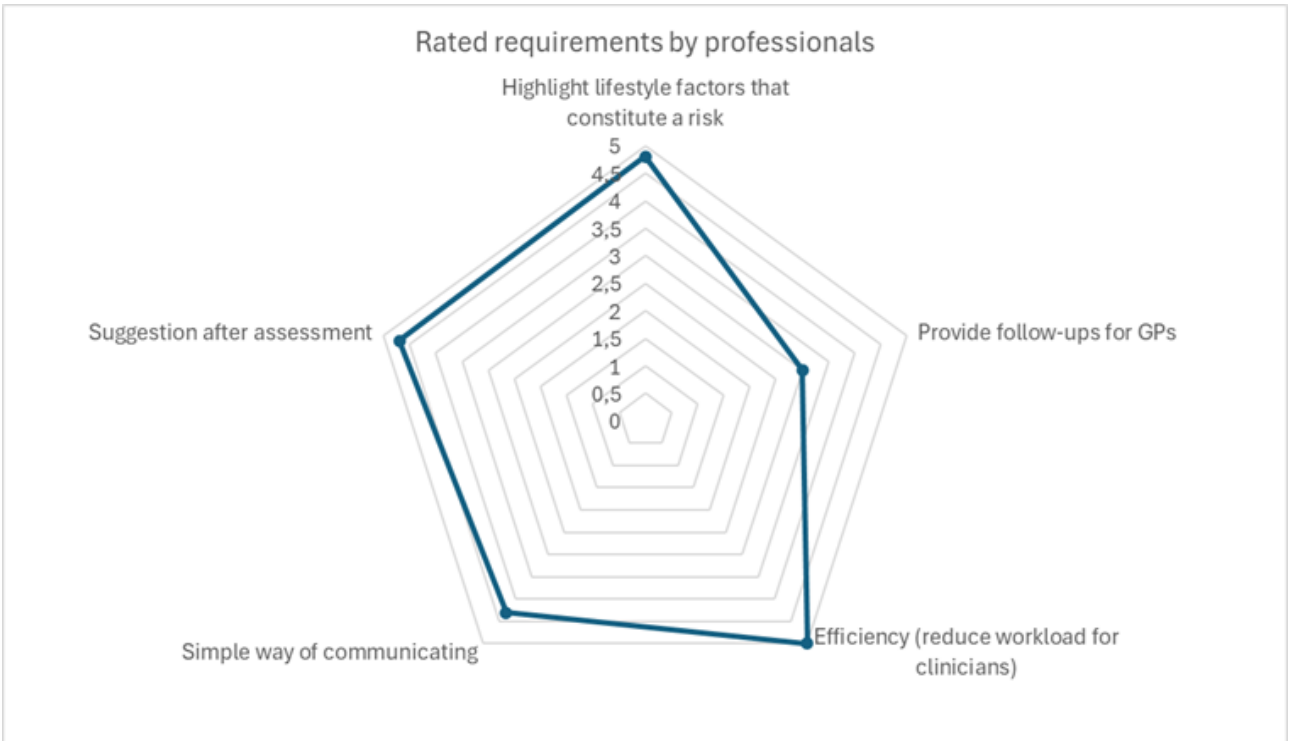
Design, Functionalities, and Evaluation of the LETHE App

Results of the Requirement Workshop and End-User Testing

The workshop together with the clinical and technical partners identified several requirements regarding the LETHE app, including the highlighting of lifestyle factors that constitute a risk or where the participant could improve. Additional observations include the LETHE app’s provision of guidelines to identify and flag study participants who are not using it. The communication between the specialists and the study participants should be simple. There was also an emphasis on data transfer and data sharing, highlighting that study participants can independently input data without professional assistance, such as for questionnaires.

The identified requirements were categorized and ranked via a survey. The main outcomes for the LETHE app are summarized in Figure 2. The highest-ranked requirements were to reduce the workload for professionals and highlight the lifestyle factors that constitute a risk. The least important requirement was to provide follow-ups for general practitioners. Those main findings were considered when designing the prototype.

Figure 2. Categorized and ranked requirements for the LETHE app. GP: general practitioner.



While conducting the user evaluation with the 4 participants at the Department of Neurology at the Medical University of Vienna, several key findings were obtained. The size of the user interface elements was criticized as being too small, and the tiles were not understood as large buttons. A section about the summary of the previous week as well as the share function of third-party apps (eg, for cognitive training or videos), which allows users to view content on another device via a QR code or by emailing it to themselves, was not considered intuitive to understand. The given tasks were missing additional information and were not clear without further explanation. Suggestions from the user testing included adding a notification for remembering to eat vegetables and fruits, as well as the marking of special input data values.

Functionalities of the LETHE App

Overview

The trial incorporates the LETHE app, an Android-native app using Java, to facilitate digital aspects and streamline data collection from study participants. The LETHE app has been released through the Google Play Store and is publicly available for download. However, access to the app's content is restricted to authorized study participants. Data storage and retrieval are managed through a .NET 7 (Core; .NET Foundation) application programming interface, which exposes data via GraphQL (Meta Platforms) protocols for external interaction with a PostgreSQL database (PostgreSQL Global Development Group). Figure 3 illustrates the modules of the LETHE app and the interaction with the LETHE CTMS.

The LETHE app correspondingly aligned the lifestyle domains with a specific digital feature based on the FINGER study [5] and included the additional domains of *sleep/relaxation* and *social activity*. Figure 4 illustrates a comparison between the in-person and digital components of the lifestyle intervention trial.

Figure 3. Overview of the modules, functionalities, and interactions between the LETHE app and the LETHE clinical trial management system (CTMS).

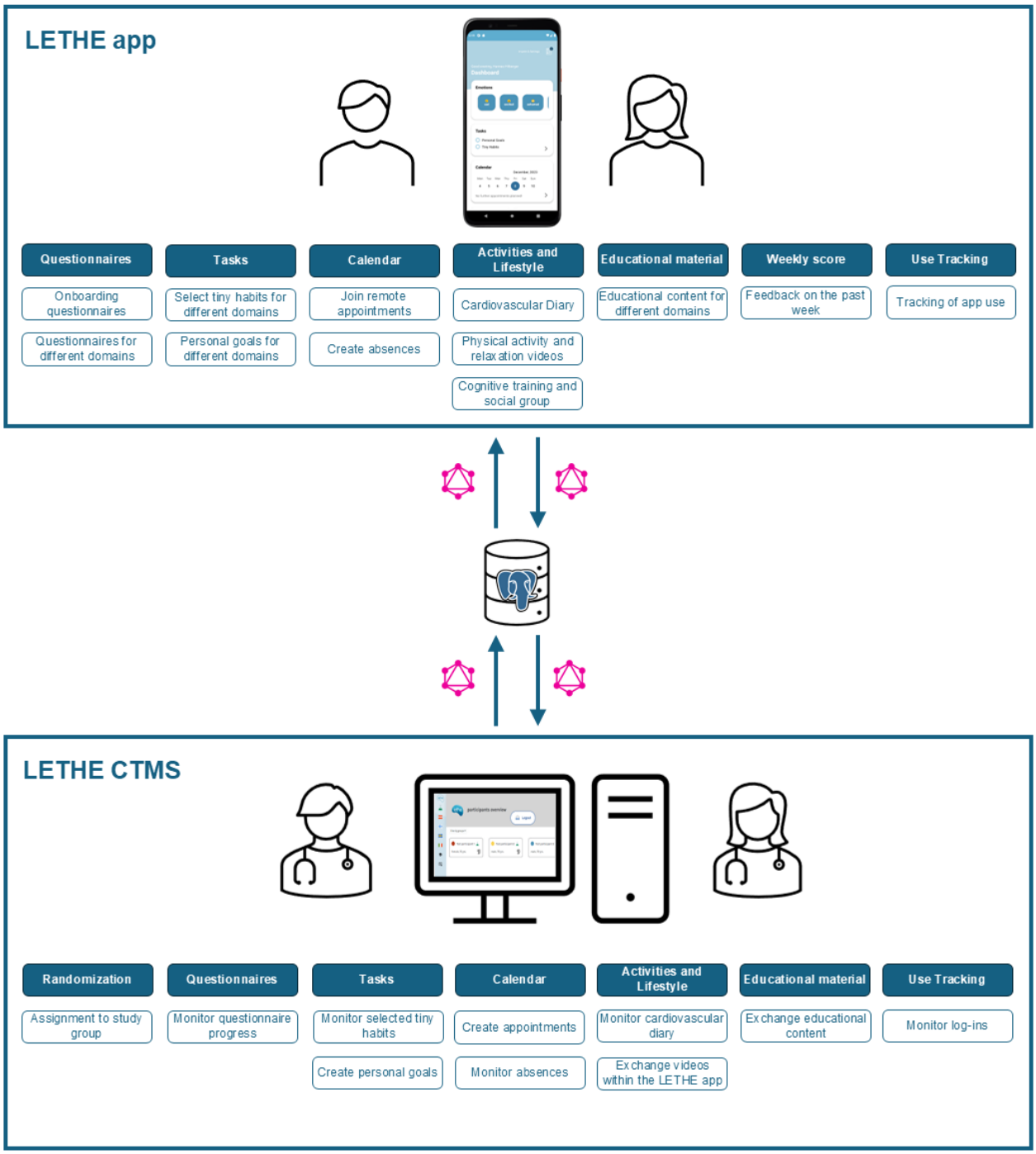
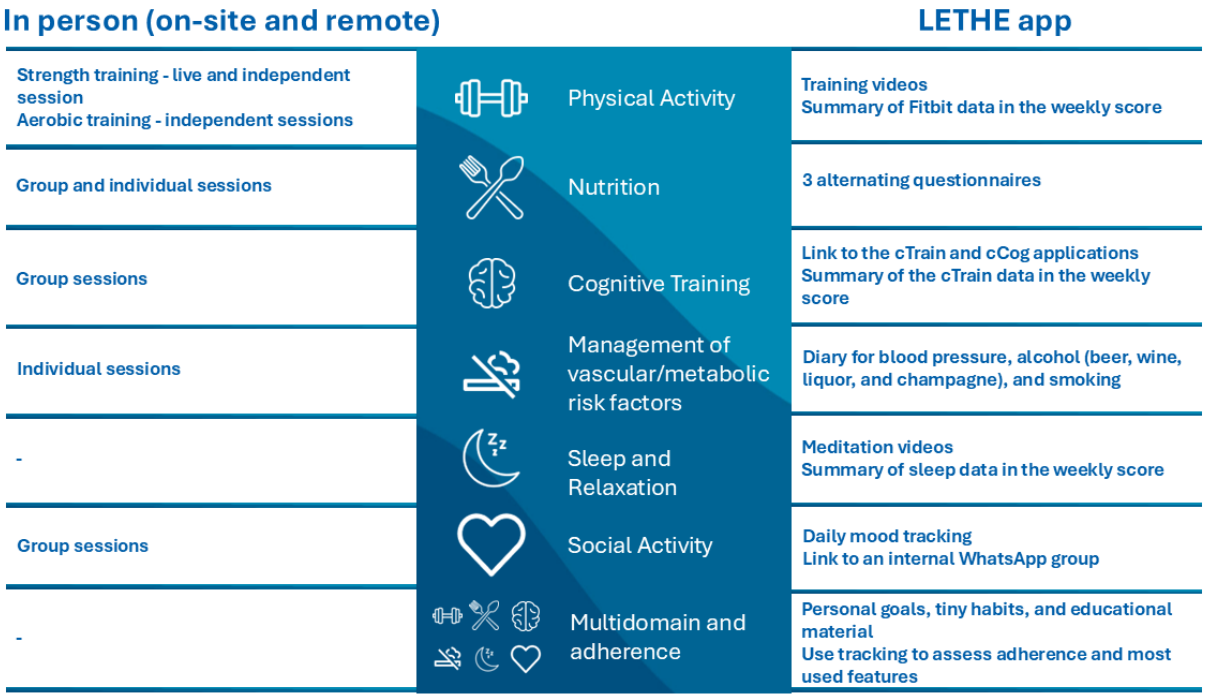


Figure 4. Comparison between in-person and digital components for each lifestyle domain.



After completing the baseline study visit and initial questionnaires through the LETHE app, participants are randomly assigned to either the intervention or control group through the LETHE CTMS. Subsequently, the LETHE app dynamically adjusts the content based on group assignment. Both groups have access to the calendar, settings, questionnaires, mood tracking, and educational content. However, the intervention group has additional functionalities, such as personalized activities and the LETHE lifestyle program.

The LETHE app supports 5 languages (English, Finnish, Italian, Swedish, and German), with translations provided by the study centers. To enhance flexibility for features such as tiny habits and questionnaires, most of the translated content is dynamically retrieved, facilitating faster translation adjustments and content extensions. The translation process involved preparing an initial draft in English, which was then translated by native speakers at the clinical centers. The translations were reviewed to ensure accuracy for each region.

Designed with an older population in mind, the LETHE app uses a tile-based approach with large tiles as entry points for each functionality. This design allows individual elements to be hidden, shown, or reused, creating a modular structure. Following this approach, each component could be packaged independently, facilitating potential reuse in other research projects with similar requirements. Screenshots of each functionality of the LETHE app can be found in Multimedia Appendix 3.

Questionnaires

The questionnaires may be completed at single or multiple time points, with variations based on study group or participant gender. Each questionnaire consists of a title, due date, progress bar, and background color divided into green, yellow, and red, indicating the time remaining to complete the questionnaire. Participants can answer questions across multiple sessions, with answers cached on the device and sent only upon completion. The technical structure adheres to the Fast Healthcare Interoperability Resources (FHIR) standard [63], ensuring interoperability with other systems and enabling the questionnaire module's reuse by incorporating different FHIR questionnaires while maintaining the same structure.

LETHE App Dashboard

Once assigned to either the intervention or control group, participants unlock access to the LETHE app dashboard, which consists of different tiles to access various functionalities. Participants can choose here how they feel daily. Each possible mood item is accompanied by a short description and an emoticon for visualization. The list of the available mood items is dynamically adaptable and could expand in future development based on participant feedback.

Tasks

The *Tasks* section encompasses the personalized intervention module, comprising 2 distinct areas: personal goals and tiny habits. Personal goals are collaboratively established with clinical professionals during visits, targeting specific lifestyle domains for personalized interventions. Personal goals follow the specific, measurable, achievable, realistic, and time-bound principle [64]. Unlike tiny habits, personal goals are not predefined, but they can be packaged as a future library of personalized objectives. These goals can be set daily, weekly, or monthly, offering frequency flexibility, including options such as bidaily or biweekly. Study participants can mark goals as completed or incomplete. Each personal goal is bound to a lifestyle domain, providing participants with an indication of their personal domains for lifestyle improvement.

Tiny habits provide practical everyday tips and behavioral suggestions to help individuals implement manageable healthy

habits [65,66] and is adapted from the StopDia library [34], which is available in Finnish under a CC BY 4.0 license [67]. Tiny habits can be individually set by participants and are available for all lifestyle domains. Each tiny habit includes a description of the activity, a health fact about its benefits, and a place where it should be performed. There are >500 tiny habits in total to choose from, and they are assigned to a lifestyle domain. Given the subjective nature of tiny habit completion, participants are asked weekly about their success based on perceived completion in a questionnaire.

Calendar

The calendar displays all study-related appointments, both remote and in person, with a monthly and daily view. Remote visits can be joined directly from the daily view, which features larger tiles for individual appointments tailored to the target group. Participants can add vacation or unavailability periods, marking them as absences in the calendar. The absence menu is distinct in contrast to all other functionalities, providing a holiday like feel for participants.

Activity and Lifestyle

This section encapsulates the comprehensive LETHE app lifestyle program, featuring various activities tailored to distinct lifestyle domains. In the "Diary" section, study participants can log different health metrics such as blood pressure, alcohol consumption, and daily cigarette intake, directly contributing to the "Management of vascular/metabolic risk factors" lifestyle domain.

A range of activities aligned with different lifestyle areas unfold in the horizontal list at the top of the screen. It starts with a variety of videos dedicated to the areas of "Physical Activity" and "Relaxation." For the "Cognitive Training" domain, there is a link to an external application called cTrain, a cognitive training game package encompassing games similar to those used in the original FINGER study [5,68].

For the "Social Activity" domain, participants from the intervention group can engage with a country-specific and internal WhatsApp group facilitated by clinical staff, ensuring exclusivity for trial participants. The last tile in the list serves as an educational hub, presenting web resources including web feeds provided by professionals via the LETHE CTMS for each lifestyle domain.

Weekly Scores

The weekly score with motivational feedback messages offers participants feedback on the previous week. The feedback message adjusts if the score increased or decreased. The score displays 3 rings representing "App Data/Lifestyle," "Fitness," and "Brain Training." The "App Data/Lifestyle" category includes how often the LETHE app is used. The "Fitness" section displays Fitbit data, including step count, sleep duration, and individual fitness sessions, whereas "Brain Training" presents an overview of cognitive training data.

Use Tracking

To obtain insight into participants' adherence and identify features of particular interest to older individuals in a lifestyle intervention app, a mechanism for use tracking was developed. Each time a participant opens the LETHE app, a unique ID is assigned for the session. Every screen is assigned to an event, and each event is saved with new time stamps when navigating through screens until the participant closes the app. This monitoring technique aims to understand useful features and incorporate measures for adherence analysis [69].

Evaluation of the LETHE App

Overview

Overall, of the 156 participants, 123 (78.8%) provided responses. The distribution of SUS scores in both the intervention group (P=.03) and the control group (P=.02) deviated significantly from normality. However, there was no significant difference in SUS scores between the groups (P=.18).

Intervention Group

A total of 78 participants from the intervention group were invited for the SUS, and 66 (85%) responses were received. Of these 66 participants, 12 (18%) did not answer all the questions. The analysis revealed that the highest satisfaction levels were associated with learning how to use the app, whereas the lowest satisfaction levels were linked to the integration of various components, as shown in Figure 5 and Multimedia Appendix 4. The integration of the various components also received the most neutral responses among the questions. The question regarding the need for assistance when using the app generated the most disagreement, whereas learning about the app before use had the smallest response rate of 85% (56/66) for all questions.

Multimedia Appendix 4 provides an additional analysis for all 4 countries. Finland achieved the highest response rate (19/20, 95% of participant responses), and Sweden achieved the lowest response rate (12/18, 67%).

As 18% (12/66) of the participants did not provide complete answers, missing values were replaced with the neutral value of 3 following the approach outlined by Lewis [70] to calculate the SUS. Figure 6 highlights the presence of 1 outlier with a total score of 0, which was not excluded.

The median SUS score was 70 (IQR 55-82), which is above the average score and qualifies the app for a grade C, as shown in Table 2.

Figure 5. System Usability Scale (SUS) responses for the LETHE app in the intervention group. Panel (A) shows positively formulated questions, whereas panel (B) highlights negatively formulated questions. Green indicates high user satisfaction, and red indicates low user satisfaction. Q: question.

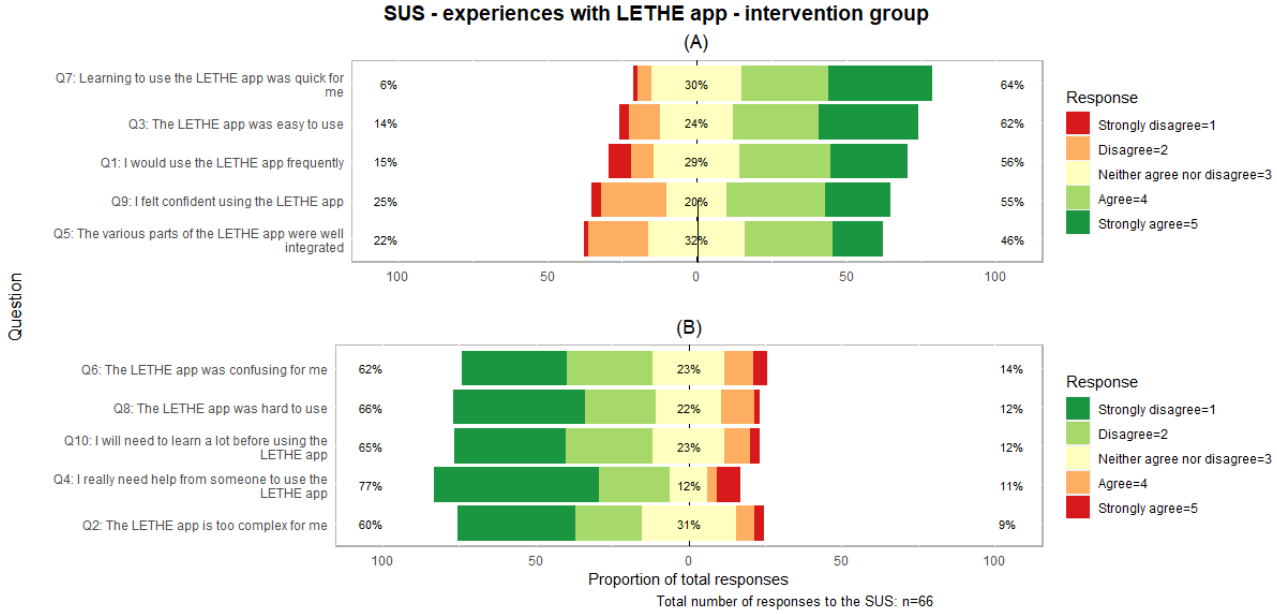


Figure 6. Box plot of the System Usability Scale (SUS) scores for each country in the intervention group.

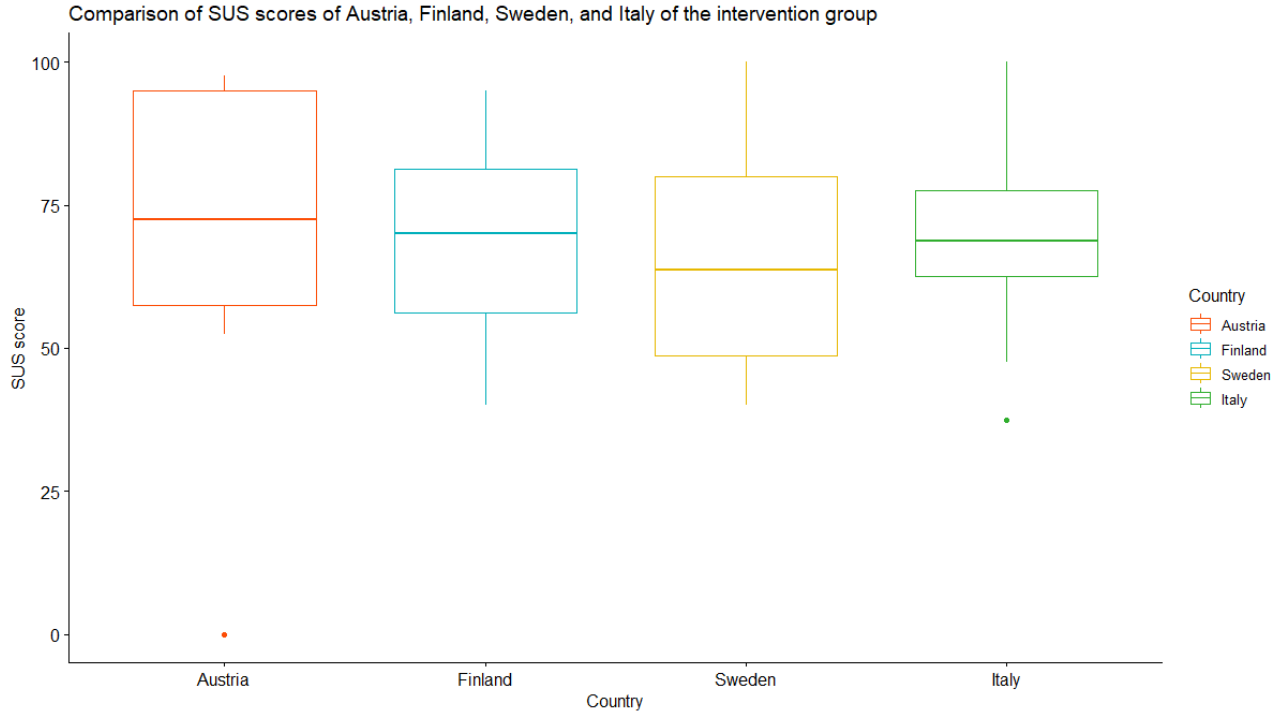


Table 2. Scores on the System Usability Scale (SUS) and its subscales for study participants in the intervention group.

	Measure[a]	Values, median (IQR)	Grade
Total (n=66)	SUS	70 (55-82)	C
Austria (n=17)	SUS	73 (58-95)	C+
Finland (n=19)	SUS	70 (57-82)	C
Sweden (n=12)	SUS	64 (49-80)	C–
Italy (n=18)	SUS	69 (63-78)	C

[a]The SUS ranges from 0 to 100, with 100 as the highest score.

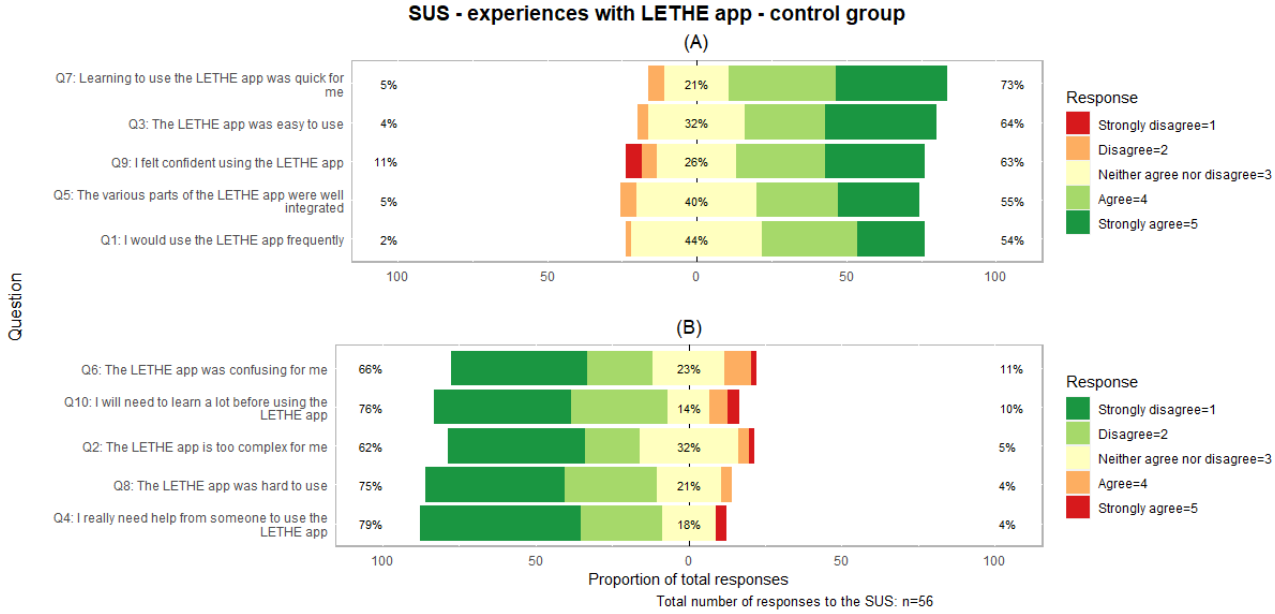
Austria achieved the highest median SUS score at 73 (IQR 58-95), whereas Sweden recorded the lowest median score at 64 (IQR 49-80). Austria was the only country where the app received a C+ score, distinguishing it from the others. Austria had 1 participant who was classified as an outlier with a total score of 0 but was still included in the analysis.

The distribution of SUS scores in Austria departed significantly from normality (P=.01). The distribution of SUS scores in Finland (P=.66), Sweden (P=.19), and Italy (P=.94) did not show evidence of nonnormality. The median SUS score did not differ significantly between the countries in the intervention group (P=.74).

Control Group

For the limited version of the LETHE app, 73% (57/78) of the study control group participants responded. Notably, 18% (10/57) of the respondents did not answer all questions. Figure 7 and Multimedia Appendix 5 indicate that the highest satisfaction was achieved in terms of how quickly participants could learn to use the LETHE app, followed by its ease of use. The only instance of a "Strongly disagree" response was regarding the confidence in using the LETHE app. The question about the LETHE app's frequent use gathered the most neutral responses, possibly due to its limited functionalities compared to the intervention version. A total of 79% (44/56) of the participants disagreed with the idea of needing help from others while using the LETHE app. Similarly to the intervention group, the question about learning about the LETHE app before using it received the smallest response rate of 98% (56/57) of all questions.

Figure 7. System Usability Scale (SUS) responses for the LETHE app in the control group. Panel (A) shows positively formulated questions, whereas panel (B) highlights negatively formulated questions. Green indicates high user satisfaction, and red indicates low user satisfaction. Q: question.



In the control group, Italy exhibited the lowest response rate, with 60% (12/20) of the participants responding, whereas Austria had the highest response rate, with 85% (17/20) of the participants responding. A detailed breakdown of responses per country is available in Multimedia Appendix 5.

To calculate the SUS, all missing values were substituted with the neutral value of 3, mirroring the approach taken with the intervention group. Table 3 illustrates that the median SUS score was 73 (IQR 63-90), corresponding to a B– grade.

Table 3. Scores on the System Usability Scale (SUS) and its subscales for study participants in the control group.

	Measure[a]	Values, median (IQR)	Grade
Total (n=57)	SUS	73 (63-90)	B–
Austria (n=17)	SUS	90 (73-95)	A+
Finland (n=16)	SUS	69 (65-76)	C
Sweden (n=12)	SUS	64 (50-84)	C–
Italy (n=12)	SUS	78 (54-90)	B+

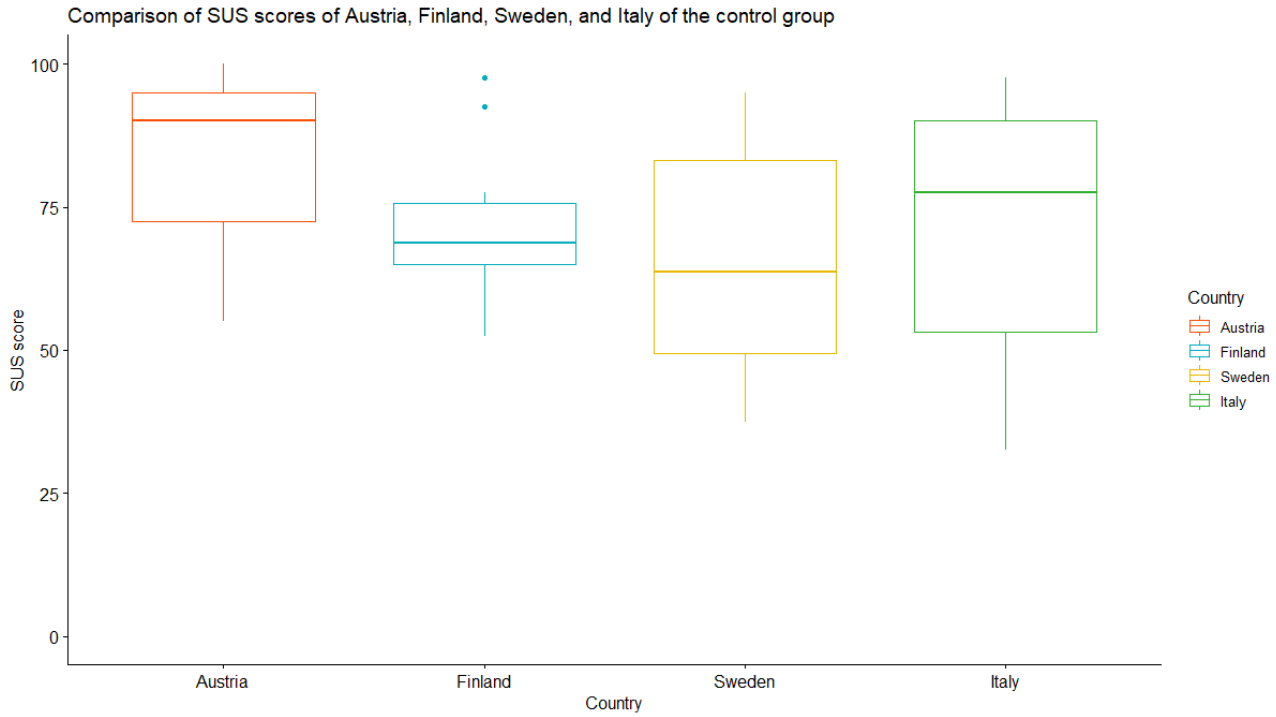
[a]The SUS ranges from 0 to 100, with 100 as the highest score.

Austria had the highest median SUS score at 90 (IQR 73-95), giving the app an A+ grade, whereas Sweden had the lowest median score at 64 (IQR 50-84), giving the app a C– grade. The app obtained a C grade in Finland and a B+ grade in Italy, with median scores of 69 (IQR 65-76) and 78 (IQR 54-90), respectively. Notably, in contrast to the intervention group, the

median SUS score was higher in Austria and Italy and lower in Finland, and it stayed the same in Sweden. Figure 8 visualizes the box plot for each country.

The SUS scores of Austria (P=.03) and Finland (P=.047) departed significantly for normality. The distribution of the SUS scores of Sweden (P=.41) and Italy (P=.11) did not show evidence of nonnormality. The SUS scores did differ between the countries in the control group (P=.03). Pairwise comparisons showed that the median SUS score of Austria was significantly different than the median SUS score of Sweden (P=.02). For the other pairwise comparisons, no significant difference was found. Specifically, the comparisons between Austria and Finland (P=.33), Austria and Italy (P=.36), Finland and Sweden (P=.99), Finland and Italy (P=.99), and Sweden and Italy (P=.99) all yielded P values greater than the threshold for statistical significance.

Figure 8. Box plot of the System Usability Scale (SUS) scores for each country in the control group.



Design, Functionalities, and Evaluation of the LETHE CTMS

Feedback From the Workshop

After an examination of the LETHE study protocol and the examples from the literature [55-57] for a design perspective, initial mock-ups for the LETHE CTMS, a web application, were created. The mock-ups were kept very general and built the foundation for the discussions in the workshop with health professionals. These mock-ups encompassed (1) an overview page featuring details on all study participants, (2) a dedicated view for an individual study participant, (3) the conceptualization of an AI simulation, and (4) a data entry page aligned with the study protocol.

The feedback on the mock-ups included the exclusion of real names given that all study participants would be collectively visible on one page in the mock-up. Emphasis was placed on the immediate visibility of adherence using colors (eg, green, yellow, and red) and dropouts. In addition, diverse roles were implemented to restrict access to sensitive information, such as ensuring that health professionals could only view data related to their own country as well as a blinded role to not see the group of the study participants. Workshop participants also expressed the desire for a descriptive overview of the dataset variables.

Subsequently, the discussion turned to the overview page, where all study participants are listed. Key considerations included the inclusion of adherence information and the need for clear differentiation between participants from the intervention and control groups. Given the sizable participant list of 156 individuals, there was a request for sorting and search functionalities.

During discussions regarding the page presenting information for an individual study participant, workshop participants advocated for distinct perspectives tailored to individuals in the intervention and control groups, emphasizing adherence and dropout summaries. Further dialogues included role-specific views, aiming for the visibility of only relevant data. The workshop participants also expressed a preference for a comprehensive presentation of lifestyle categories, including detailed information and distinct sections for different clinical professionals as well as additional risk factors such as the consumption of alcohol or cigarettes. Requests for additional functionalities included the presentation of the Clinical Dementia Rating [71], allergies, questionnaire responses and scores, notes on participant interactions, details of adverse events such as muscle pain after a prescribed aerobic workout, and automated calculations such as BMI. Another necessity identified was the inclusion of a data export functionality.

When considering the AI dementia risk simulation, there was a discussion about whether it could be used during the study intervention or solely for research purposes. The AI risk simulation is currently available only to researchers due to ethical considerations. As the models used in the simulation have not yet been fully tested or validated, they were not implemented with real participants at this stage. Moreover, the risk information generated by the simulation is not disclosed to participants. Should the disclosure of this risk information be considered in the future, it would require careful evaluation. This process would involve consulting with both the participants and experts in ethics to assess the potential psychological impacts and the broader ethical implications of sharing risk-related information with participants.

The workshop discussed the insertion of all CRF data into the LETHE CTMS, encompassing visit-related data and scores such as the Clinical Dementia Rating and Neuropsychological Test Battery as in the FINGER study. Workshop participants underscored the inclusion of medications, validation checks, and prefillable fields. The study protocol served as a reference for defining the data entry page fields.

Following the conclusion of the workshop, participants were given a clickable prototype. Subsequently, a follow-up workshop was conducted to incorporate adaptations based on the initial workshop insights. The most significant modification was made to the data entry page, necessitating a complete restructuring due to a higher-than-anticipated number of data fields. After several iterations, a design for the beginning of the study was proposed.

Functionalities of the LETHE CTMS

Overview

Following the completion of the LETHE CTMS design, pivotal components were identified to commence the clinical trial. This included an overview of all study participants, a detailed view of a study participant, and the data input portal for clinical professionals. The LETHE CTMS was developed as a web application using React and uses the same backend infrastructure as the LETHE app. The LETHE CTMS is deployed as a Docker container (Docker, Inc). Once a new version is pushed to the main branch of the GitLab repository, a Continuous Integration/Continuous Delivery or Deployment (CI/CD) pipeline is automatically triggered, handling the build, release, and deployment process. After the container is pushed to the GitLab container registry, deployment is carried out using an Ansible playbook (Ansible Inc).

The LETHE CTMS has a consistent design throughout the application, achieved by using a fundamental page structure featuring headers, card-based content representation, a uniform color scheme inspired by the LETHE project colors, and font selection. Ensuring that the data are relevant only to the relevant users was achieved through ongoing communication with end users, and various roles were defined. There are roles that have access to all participants from all study centers; roles that are blinded and, therefore, can only enter visit forms; and roles that are not blinded and can have access to all participant details from their study center.

The LETHE CTMS underwent iterative modifications over time driven by continuous user feedback and evolving requirements. In particular, each section has been given a standardized nomenclature. At the moment, the following sections are included: overview page, detail page, clinician data entry page (containing the electronic CRF [eCRF]), and the configuration pages to adapt the content in the LETHE app. For now, the AI risk simulation is only available for research purposes and not visible to the end users, so a further description is not provided. Screenshots of the different pages can be found in Multimedia Appendix 6.

Overview Page

The overview page serves as a visual representation of all study participants within the selected country. Upon initializing this page, the study participants are loaded based on the study site affiliation of the clinical professional. The header features a legend outlining variables encompassing distinct adherence pathways and the respective participant groups. Color-coded adherence pathways range from “low” (red) to “medium” (yellow) to “high” (green) and include a category of “not calculated.” These pathways are grouped into distinct types, with more details provided on the detail page.

The central section of the page displays all participants from the current site within individual rounded boxes, which encompasses variables such as the participant’s ID, and other essential participant information, including adherence, group affiliation, age, and gender, is also presented.

There is also the possibility to access the configuration pages, allowing for the modification of content related to educational material and videos on the LETHE app.

Detail Page

Upon selecting an individual on the overview page, a more detailed view is presented, encompassing comprehensive data on that particular participant.

The detailed view interface is organized into 3 columns, each with functional components. In the first column, baseline visit data are featured, encompassing demographics, cardiovascular metrics, comorbidities, medications, allergies, and blood values. Users can review completed and pending questionnaires as well as access the responses and corresponding scores. This column concludes with a messaging feature that allows professionals to schedule notifications directly sent to study participants’ smartphones.

The second column commences with notes about adverse events, the entering and tracking of personal goals, and the monitoring of the selected tiny habits and cardiovascular risk factor data entered into the LETHE app. Professionals can also view participant absences to adjust appointment scheduling and reschedule personal goals accordingly.

In the last column, professionals can schedule recurring remote or in-person meetings with study participants together with reminders. Each meeting can be marked as completed to help track the adherence. Participant-professional contacts are recorded, and additional information is displayed, including last

LETHE app log-in, participant consent status, and participation in substudies.

At the bottom of the page, adherence to various categories, such as app use, cognitive activity, diet, and physical activity, is visually displayed, providing an overview of the participant's engagement with key aspects of the study protocol. Table 4 shows the thresholds used to assign each pathway to one of the adherence levels. On the basis of adherence and lifestyle domain, participants receive tailored messages on specific weekdays: physical activity on Mondays, app use on Tuesdays, diet on Wednesdays, cognitive activity on Thursdays, cardiovascular risk factors and social interaction on Fridays, and relaxation and sleep on Saturdays. Participants consistently in the red pathway receive domain-specific or holistic messages. In addition, tiny habit messages are sent on Mondays, Wednesdays, and Thursdays based on reported engagement.

Table 4. Adherence pathways and thresholds for different categories.

	Data type	Instrument used to measure the data and frequency of measurement	Green threshold	Yellow threshold	Red threshold
Physical activity	Activity min and frequency (moderate to vigorous intensity)	Fitbit; weekly (Mondays); passive	Months 1-3: 30-45 min (1 time per wk); months 4-6: 30-45 min (2-3 times per wk); months 7-9: 30-60 min (3-4 times per wk); months 10-24: 45-60 min (3-5 times per wk)	Months 1-3: 15-29 min (1 time per wk); months 4-6: 15-29 min (1 time per wk); months 7-9: 15-29 min (1-2 times per wk); months 10-24: 15-44 min (1-2 times per wk)	1 wk with 0 min or 3 weeks in a row in the yellow path
App use	Changes between the screens within the LETHE app	App; weekly (Tuesdays); passive	≥3 screens (4-7 d)	≥3 screens (0-3 d)	0 screens for 14 d
Diet patterns	Completing food-monitoring item questionnaire within the LETHE app (3 alternating questionnaires)	App; weekly (Wednesdays); active	2/3 or 3/3 blocks completed (in 3 wks)	1/3 blocks completed (in 3 wks)	0% completed (in 3 wks)
Cognitive training	Use of cognitive training program	cTrain; weekly (Thursdays); active	2-3 times per wk	0-1 times per wk	0 times for 3 wk

Data Entry Page

Diverging from the workshop, a series of design iterations led to the decision that the data entry page would act as a repository for information gathered from all study visits, whereas tasks such as contacting study participants and scheduling appointments were moved to the detailed view of an individual participant.

Within the eCRF, each primary section, covering screening, baseline, and subsequent visits, contains subforms that address different documentation areas (eg, forms to report results for blood tests, medication use, and neurological assessments such as the Mini-Mental State Examination [72]).

Each subform within the eCRF incorporates features such as automatic calculations and validation checks, whereby abnormal values trigger a visual highlighted in yellow. This ensures data accuracy and enhances the efficiency of data entry. The eCRF design, embedded with these functionalities, serves as a comprehensive and streamlined platform for recording and managing diverse aspects of study visit documentation.

Evaluation of the LETHE CTMS

Overview

This section analyzes the feedback from the CTMS survey completed by health professionals. First, user-specific details, such as their role in the project and within the application are presented. Second, the quantitative and qualitative feedback for each of the functionalities and the design is analyzed. Finally, a ranking of the state-of-the-art features is presented. The survey refers to the LETHE CTMS as *LETHE dashboard*, but both terms are interchangeable.

Users of the LETHE CTMS

The survey targeted all users of the LETHE CTMS, gathering general information such as their role in the project, their function within the application, and the frequency of their application access. When the survey was conducted, 21 users had access to the LETHE CTMS, and 15 (71%) responded to the survey. The summarized results are presented in Table 5. Most respondents held roles as coordinators or in unblinded positions, with access to almost all features. Of the 15 respondents, 12 (80%) accessed it at least once a week.

Table 5. Role of professionals in the LETHE project and in the LETHE clinical trial management system (CTMS) and frequency of CTMS use (N=15).

Answer option	Professionals, n (%)
"Could you please describe your role in the project?"	
Coordinator	1 (7)
Digital coach	2 (13)
Neuropsychologist	3 (20)
Nutritionist	2 (13)
Physiotherapist	1 (7)
Principal investigator	1 (7)
Study nurse	4 (27)
Study physician	1 (7)
"What dashboard role do you have?"	
Blinded role	1 (7)
Coordinator role	6 (40)
Unblinded role	8 (53)
"How frequently do you access the dashboard?"	
Daily	1 (7)
More than once a week	6 (40)
Once a week	5 (33)
More than once a month	1 (7)
Once a month	2 (13)

Feedback on the Design and Functionalities of the LETHE CTMS

Figure 9 and Table 6 indicate that the highest level of user satisfaction was related to the comprehensive display of relevant participant information in an easily understandable way, with a score of 100% above neutral. A detailed breakdown of all quantitative responses can be found in Multimedia Appendix 7. Overall, feedback was positive, reflecting successful integration of design approaches from workshops and feedback. Users found the LETHE CTMS intuitive, with a visually appealing and clear layout featuring consistent colors, logically divided sections, and pictograms. Pop-ups during item saving aided navigation and provided feedback on actions. Suggested improvements included reducing the steps to move between participants, addressing inconsistencies in single- and double-click requirements, and ensuring visibility of study participant phone information on the participant overview page.

Figure 9. Survey responses for the LETHE clinical trial management system from health professionals. Green indicates high user satisfaction, and red indicates low user satisfaction.

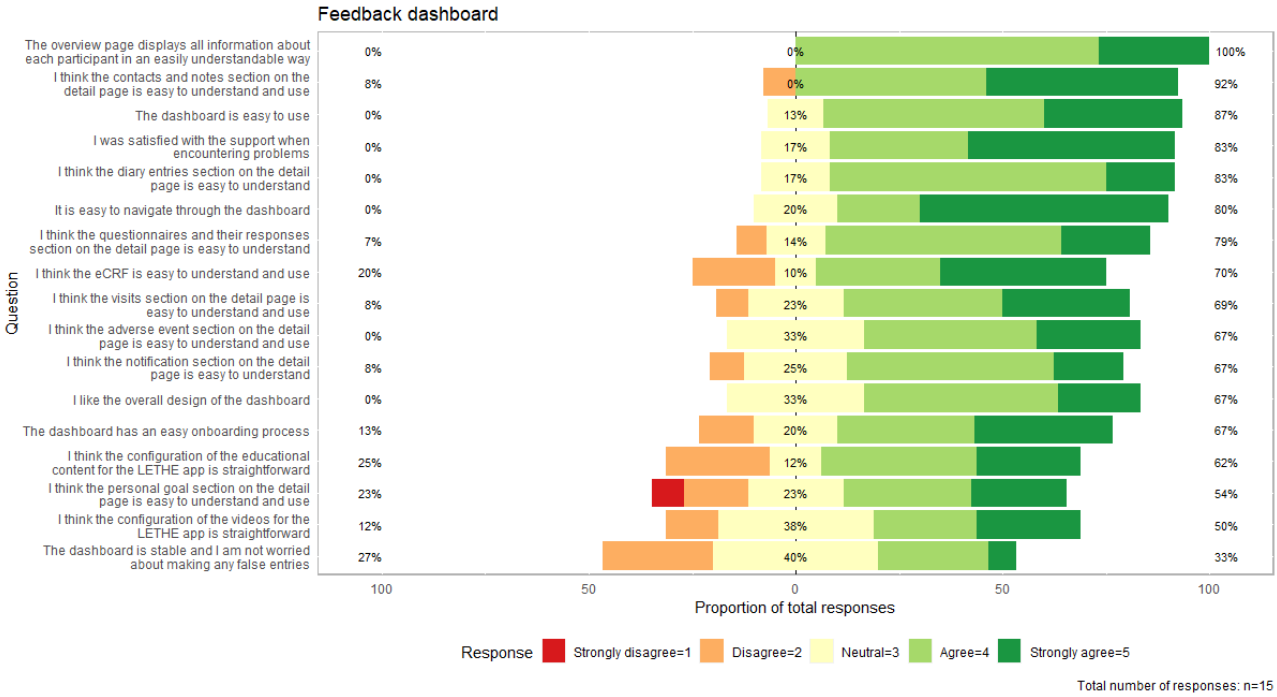


Table 6. Means and SDs of the LETHE clinical trial management system feedback survey targeting the different functionalities.

	Values, mean (SD)[a]
“The overview page displays all relevant information about each participant in an easily understandable way” (n=15)	4.27 (0.46)
“I think the Contacts/Notes Section on the Detail Page is easy to understand use” (n=13)	4.31 (0.85)
“The Dashboard is easy to use” (n=15)	4.20 (0.68)
“I was satisfied with the support when encountering problems” (n=12)	4.33 (0.78)
“I think the Diary Entries Section on the Detail Page is easy to understand” (n=12)	4.00 (0.60)
“It is easy to navigate through the Dashboard” (n=15)	4.40 (0.83)
“I think the Questionnaires and their Responses Section on the Detail Page is easy to understand” (n=14)	3.93 (0.83)
“I think the eCRF is easy to understand and use” (n=10)	3.90 (1.20)
“I think the Visits Section on the Detail Page is easy to understand and use” (n=13)	3.92 (0.95)
“I think the Adverse Event Section on the Detail Page is easy to understand and use” (n=12)	3.92 (0.79)
“The Dashboard has an easy onboarding process” (n=15)	3.87 (1.06)
“I like the overall design of the Dashboard” (n=15)	3.87 (0.74)
“I think the Notification Section on the Detail Page is easy to understand” (n=12)	3.75 (0.87)
“I think the configuration of the Educational Content for the LETHE App is straightforward” (n=8)	3.63 (1.19)
“I think the Personal Goal Section on the Detail Page is easy to understand and use” (n=13)	3.46 (1.27)
“I think the configuration of the Videos for the LETHE App is straightforward” (n=8)	3.63 (1.06)
“The Dashboard is stable and I am not worried about making any false entries” (n=15)	3.13 (0.92)

[a]The scale ranges from 1 (*Strongly disagree*) to 5 (*Strongly agree*). A higher level of agreement indicates greater satisfaction.

Concerning the onboarding process, users expressed that it involved multiple steps and should be streamlined further. However, once granted access, the general opinion was that the process was straightforward. The lowest satisfaction regarding functionalities was observed in the educational content and video configuration, where users missed options for uploading videos and materials instead of web links. The stability and prevention of false entries, particularly in the eCRF, received the lowest satisfaction rating. Users reported issues with scrolling affecting eCRF values and occasional problems with

saving values. Suggestions included implementing a log to track data edits for added transparency, a measure already in place but not visible to users. Another notable disagreement in the *Personal Goals* section was that users found it difficult to set goals at regular intervals, and the lifestyle domain associated with each personal goals was not visible in the detailed view of a study participant.

Feature suggestions included displaying follow-up data, editing values in different sections, providing additional Fitbit data insights, and summarizing clinical test names and scores at the bottom of the eCRF page. Positive aspects highlighted satisfaction with support, emphasizing quick response times, correction of implementations, and clear explanations.

Other feedback suggested the ability to send messages to groups of study participants as well as create group meetings. Overall, those surveyed commented on the effectiveness of the LETHE CTMS for study staff and participants, noting a continuous improvement since its first use.

Ranking the Most Useful and Beyond-State-of-the-Art Features

One of the survey questions asked respondents to rank various features integrated into the LETHE CTMS based on their perceived usefulness and how they surpassed state-of-the-art capabilities. As shown in Table 7, the integration of personalized activities (mean 2.23, SD 1.17) and real-time appointment planning (mean 2.46, SD 1.51) stood out as the most impactful features. An additional comment from a respondent highlighted the novelty of the interplay between the LETHE CTMS and the LETHE app, with a note that further optimization was needed. In addition, the LETHE CTMS holds the potential to serve as both an eCRF and a tool to assist in intervention delivery. This dual functionality was seen as a valuable development for the future to meet the evolving needs of professionals conducting clinical trials.

Table 7. Ranking of different features based on surpassing state-of-the-art capabilities, with R1 being the best-ranked feature and R7 being the worst-ranked feature.

	Choices (n=13), n (%)							Total (n=91), n (%)
	C1[a]	C2[b]	C3[c]	C4[d]	C5[e]	C6[f]	C7[g]	
R1	2 (15)	2 (15)	1 (8)	0 (0)	1 (8)	4 (31)	3 (23)	13 (14)
R2	1 (8)	0 (0)	0 (0)	0 (0)	1 (8)	4 (31)	7 (54)	13 (14)
R3	1 (8)	0 (0)	2 (15)	4 (31)	3 (23)	2 (15)	1 (8)	13 (14)
R4	0 (0)	3 (23)	3 (23)	3 (23)	1 (8)	2 (15)	1 (8)	13 (14)
R5	1 (8)	1 (8)	3 (23)	2 (15)	5 (38)	0 (0)	1 (8)	13 (14)
R6	2 (15)	5 (38)	1 (8)	3 (23)	1 (8)	1 (8)	0 (0)	13 (14)
R7	6 (46)	2 (15)	3 (23)	1 (8)	1 (8)	0 (0)	0 (0)	13 (14)
Total (n=91)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	13 (14)	91 (100)

[a]Overview page of different study participants in different countries, including their digital intervention adherence pathways (mean 5.08, SD 2.43).

[b]Direct notifications to study participants via the dashboard (mean 4.84, SD 1.99).

[c]Electronic case report form (mean 4.69, SD 1.80).

[d]Immediate results of onboarding and questionnaires on the dashboard (mean 4.54, SD 1.40).

[e]Configuration of the content of the app via the dashboard (mean 4.15, SD 1.68).

[f]Real-time planning of appointments within the dashboard for the app (mean 2.46, SD 1.51).

[g]Integration of personalized intervention tasks (*Personal Goals*; mean 2.23, SD 1.17).

Discussion

Principal Findings

This paper outlines the design process, functionalities, and evaluation of the digital intervention study components within the LETHE project, focusing on older individuals at risk of cognitive decline and clinical professionals. The design process involved multiple sessions with potential end users and clinical experts. Furthermore, the setup is currently being evaluated in a 2-year intervention study, and the first results have been presented [11,69].

The original FINGER multidomain trial, along with related studies, demonstrated that cognitive benefits can be achieved through a 2-year intervention. As a result, this duration was established as the timeline for the LETHE trial. However, there is only limited information available on longer studies supported by ICT components, such as digital apps requiring daily interactions and wearable devices.

These requirements made it necessary to carefully design the study in terms of human support and create user-friendly ICT components for both participants and clinical professionals. The learnings regarding ICT use in clinical studies, as well as data collected in intervention studies (clinical data and digital biomarkers), are essential and can influence future studies and clinical trial setups using ICT.

To emphasize the importance of user-friendly design, this paper presents the process from requirement gathering to the design of the components. The findings will be evaluated and will inform further improvements to the setup.

Overall, the median SUS score of the intervention group of 70 (IQR 55-82) was comparable to that of the control group of 73 (IQR 63-90). The slightly higher median score of the control group might be attributed to the less complex functionalities compared to those of the intervention group version, specifically the absence of personalized and intervention activities. However, both median scores surpassed 68, indicating above-average user acceptance and satisfaction according to SUS guidelines [62]. Given that the initial user experience for the control group was quite positive even with fewer features, there is hope that they will remain engaged throughout the trial. It is anticipated that these values will improve as participants gain more experience with the LETHE app during the 2-year RCT [73]. Evaluating the other time points will give us more information about the longer-term usability and user engagement.

Noteworthy is the variation in median SUS scores between countries, with Austrian participants providing the highest scores (intervention group: median 73, IQR 58-95; control group: median 90, IQR 73-95) and Sweden providing the lowest scores (intervention group: median 64, IQR 49-80; control group: median 64, IQR 50-84). Performing a pairwise comparison using the Dunn test, a significant difference in median SUS scores between Austria and Sweden was found (P=.02).

While the LETHE lifestyle intervention program is centrally coordinated, and the activities in the 4 countries are harmonized to ensure comparable content, certain local adaptations are allowed to optimize feasibility (eg, in the detailed intervention delivery [balance between in-person and digital sessions] and how the digital components are leveraged). A qualitative study is currently underway to collect additional data on the barriers to and facilitators of the use of digital tools among participants in the LETHE trial.

Overall, users expressed satisfaction with the LETHE CTMS, and workshops on design, usability, and functionality proved beneficial. However, clarity regarding the purpose of an AI risk simulation and addressing user concerns about inaccurate data entry emerged as key challenges during the design process. Further exploration into user satisfaction with the intervention's introduction by clinical professionals across different countries is warranted for future research.

Research should investigate reasons behind decreased personal engagement in long-term ICT-supported studies. The LETHE protocol, with its 2-year ICT-supported design, offers a unique contribution to support cognitive health in older individuals.

In terms of future research, several key directions should be explored. First, a thorough evaluation of the platform's use and usability after the completion of the full 2-year trial would provide valuable insights into its long-term effectiveness. Second, further investigations could focus on the AI-driven risk simulation within the CTMS, with a particular emphasis on its potential role in guiding future interventions. Finally, exploring the adaptability of the LETHE platform among populations with low digital literacy and assessing the feasibility of a hybrid approach would be an important area for future research.

We are confident that LETHE app's diverse features, together with high retention, positive feedback and frequent use [11,69], provide promising directions for future hybrid multimodal interventions.

Strengths and Limitations

This study has several strengths, particularly in how the LETHE app and CTMS incorporate perspectives from both northern and southern Europe. This allows for a broader range of views to be considered. Moreover, this study is regularly improved through feedback from professionals, suggestions from study participants, and input from advisory board members, ensuring ongoing refinement and relevance. One limitation lies in assuming that all participants have sufficient digital skills and internet access, which was an inclusion criterion for the RCT. However, according to a recent study, the gap in digital health trends between younger and older people may vanish in 10 years as today's individuals aged <65 years are highly adapted to digital solutions [74].

Furthermore, all participants were provided with introductory materials and manuals for the LETHE app to facilitate use. Those factors could impact their ratings. Moreover, the subjective nature of the SUS, which relies on participants' self-reported perceptions of usability, might not fully grasp the nuanced usability challenges, especially among older users who may have varying cognitive and physical abilities. Another limitation is that the LETHE app content (eg, educational materials or videos) varies between countries and may influence the ratings.

Another consideration is that the initial user testing involved a small sample size, with only 4 German-speaking participants for the first user testing. However, the evaluation of the LETHE app is part of a study in which health professionals and 156 study participants are included.

The LETHE app was developed exclusively for the Android platform. This decision was based on practical considerations, including resource constraints, time limitations, and the need for a streamlined pilot phase. Focusing on a single operating system allowed for a more efficient development, testing, and release process, ensuring that the feasibility trial could be completed within the available time frame. Android was chosen due to its widespread use and compatibility with various devices, facilitating broader accessibility. To mitigate potential accessibility barriers, participants who did not have an Android device were given the option to receive one for the study. This approach enabled a standardized evaluation of the app's functionality and user experience. Future iterations of the app may consider cross-platform availability, including an iOS or web version.

Regarding the LETHE CTMS, a limitation of the survey is that most users held roles with the highest level of privileges in terms of functionalities, with frequent role changes occurring throughout the project. Improved planning of roles is advisable for future applications to mitigate this issue.

Comparison With Prior Work

Previous studies indicate declining user engagement with mHealth interventions over time due to multiple factors, such as unengaging content, complex protocols, or poor design. For instance, studies such as predictive model-based decision support for diabetes patient empowerment (POWER2DM) [75] have investigated ICT-supported health systems integrating features such as shared decision-making, personal goal setting, mood tracking, and exercise components, which is similar to the interaction between the LETHE app and LETHE CTMS but within the context of diabetes self-management.

Compared to existing digital interventions for dementia prevention, such as PRODEMOS [15] and Maintain Your Brain [16], LETHE introduces several additions and refinements. LETHE builds on standard goal setting but also includes the tiny habits method as an add-on. It also covers additional lifestyle domains, including social engagement and sleep, which are not included in PRODEMOS but are, meanwhile, part of the Lancet Commission's report [2].

In terms of data collection, LETHE uses Fitbit devices to passively monitor physical activity and sleep, whereas PRODEMOS relies on self-reported step counts. Both interventions use a hybrid approach, but LETHE's study duration is extended to 24 months (compared to 18 months in PRODEMOS) and builds on the successful FINGER protocol [5].

LETHE differs from Maintain Your Brain by offering immediate access to all modules rather than introducing them sequentially. Although LETHE does not include a separate mental health module, it collects related information through questionnaires. Maintain Your Brain, compared to LETHE, is entirely internet based and spans 3 years, with only quarterly booster activities after the first year and yearly follow-up assessments for the remaining 2 years.

Finally, unlike both PRODEMOS and Maintain Your Brain, a primary outcome of LETHE is the feasibility of a digitally supported multimodal lifestyle intervention, assessed through participant engagement with the LETHE app and a Fitbit device, attendance to study activities, and implementation of a hybrid approach that combines active and passive data collection across 4 countries.

One significant factor contributing to the unsatisfactory use of ICT components in ICT intervention studies is the usability of the functions provided. This paper provides a comprehensive account of the design process for both the app and the CTMS, aiming to maximize their usability. Subsequent publications will detail the outcomes concerning use and user satisfaction. Identified gaps in involving older adults in the design process of digital health technologies shed light on issues such as the exclusion of individuals with low digital skills during recruitment, emphasizing the importance of using diverse methods such as focus groups, interviews, or workshops [76]. While both the LETHE app and LETHE CTMS design processes incorporated various methods to solicit continuous feedback from diverse user populations, they did not explicitly include persons with low digital skills.

Conclusions

In summary, this paper presented and evaluated the design and functionalities of a comprehensive clinical trial system involving a mobile app tailored for older individuals at risk of cognitive decline. The LETHE app and a web-based monitoring and configuration system for clinical professionals, the LETHE CTMS, were thoroughly evaluated through survey-based assessments. Overall, our approach facilitates real-time interaction, providing 2 distinct applications for professionals and study participants. The results elucidated the design principles, stakeholder involvement, and essential functionalities of such an eHealth system. Subsequent research will delve into the posttrial use of the LETHE app over the 2 years, shedding further light on its effectiveness and user engagement.

Acknowledgments

The LETHE project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 101017405 [77] (accessed on January 31, 2024). FM is also supported by an ALF (Region Stockholm) grant. This research is supported by a larger consortium, including FH JOANNEUM GmbH, Medical University of Vienna, University of Perugia, Karolinska Institute, Finnish Institute for Health and Welfare, Alzheimer Europe, Foundation for Research and Technology–Hellas, Maastricht University, Kaasa solution GmbH, Innovation2Grow, the EGI Foundation, Extra Red SRL, Infotrend Innovations Company Limited, Combinostics Oy, and the Lisbon Council for Economic Competitiveness asbl.

Conflicts of Interest

None declared.

Multimedia Appendix 1
LETHE app survey.
[PDF File (Adobe PDF File), 200 KB - aging_v8i1e66660_app1.pdf]

Multimedia Appendix 2
LETHE clinical trial management system survey.
[PDF File (Adobe PDF File), 366 KB - aging_v8i1e66660_app2.pdf]

Multimedia Appendix 3
LETHE app screenshots.
[PPTX File , 25265 KB - aging_v8i1e66660_app3.pptx]

Multimedia Appendix 4
LETHE app evaluation—intervention group.
[XLSX File (Microsoft Excel File), 15 KB - aging_v8i1e66660_app4.xlsx]

Multimedia Appendix 5
LETHE app evaluation—control group.
[XLSX File (Microsoft Excel File), 15 KB - aging_v8i1e66660_app5.xlsx]

Multimedia Appendix 6
LETHE clinical trial management system screenshots.
[PPTX File , 549 KB - aging_v8i1e66660_app6.pptx]

Multimedia Appendix 7
LETHE clinical trial management system cross-table.
[XLSX File (Microsoft Excel File), 12 KB - aging_v8i1e66660_app7.xlsx]

References

1. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
2. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
3. Kivipelto M, Mangialasche F, Anstey KJ. Pivotal points in the science of dementia risk reduction. Lancet 2024 Aug 10;404(10452):501-503. [doi: 10.1016/S0140-6736(24)01546-0] [Medline: 39096928]
4. Noach S, Witteman B, Boss HM, Janse A. Effects of multidomain lifestyle interventions on cognitive decline and Alzheimer's disease prevention: a literature review and future recommendations. Cereb Circ Cogn Behav 2023;4:100166 [FREE Full text] [doi: 10.1016/j.cccb.2023.100166] [Medline: 37215433]
5. Ngandu T, Lehtisalo J, Solomon A, Levälahti E, Ahtiluoto S, Antikainen R, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 06;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
6. Thunborg C, Wang R, Rosenberg A, Sindi S, Andersen P, Andrieu S, et al. Integrating a multimodal lifestyle intervention with medical food in prodromal Alzheimer's disease: the MIND-AD randomized controlled trial. Alzheimers Res Ther 2024 May 30;16(1):118 [FREE Full text] [doi: 10.1186/s13195-024-01468-x] [Medline: 38812047]
7. Whitfield T, McConnell B, Renouf P, Mansour H, Zabihi S, Aguirre E, et al. The effect of remotely delivered lifestyle interventions on cognition in older adults without dementia: a systematic review and meta-analysis. Ageing Res Rev 2021 Dec;72:101505 [FREE Full text] [doi: 10.1016/j.arr.2021.101505] [Medline: 34757173]
8. Richard E, Moll van Charante EP, Hoevenaar-Blom MP, Coley N, Barbera M, van der Groep A, et al. Healthy ageing through internet counselling in the elderly (HATICE): a multinational, randomised controlled trial. Lancet Digit Health 2019 Dec;1(8):e424-e434 [FREE Full text] [doi: 10.1016/S2589-7500(19)30153-0] [Medline: 33323224]
9. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:9 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
10. Hanke S, Mangialasche F, Bödenler M, Neumayer B, Ngandu T, Mecocci P, et al. AI-based predictive modelling of the onset and progression of dementia. Smart Cities 2022 May 20;5(2):700-714. [doi: 10.3390/smartcities5020036]
11. Rosenberg A, Untersteiner H, Guazzarini AG, Bödenler M, Bruinsma J, Buchgraber-Schnalzer B, on behaf of the LETHE Consortium. A digitally supported multimodal lifestyle program to promote brain health among older adults (the LETHE randomized controlled feasibility trial): study design, progress, and first results. Alzheimers Res Ther 2024 Nov 21;16(1):252 [FREE Full text] [doi: 10.1186/s13195-024-01615-4] [Medline: 39574193]
12. Bruinsma J, Visser LN, Abaci A, Rosenberg A, Diaz A, Hanke S, et al. Social activities in multidomain dementia prevention interventions: insights from practice and a blueprint for the future. Front Psychiatry 2024 May 20;15:1386688 [FREE Full text] [doi: 10.3389/fpsyt.2024.1386688] [Medline: 38832328]

13. Di Lorito C, Bosco A, Rai H, Craven M, McNally D, Todd C, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and mild cognitive impairment. Int J Geriatr Psychiatry 2022 Jun 19;37(6):5730 [FREE Full text] [doi: 10.1002/gps.5730] [Medline: 35588315]
14. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
15. Eggink E, Hafdi M, Hoevenaar-Blom MP, Song M, Andrieu S, Barnes LE, et al. Prevention of dementia using mobile phone applications (PRODEMOS): protocol for an international randomised controlled trial. BMJ Open 2021 Jun 09;11(6):e049762 [FREE Full text] [doi: 10.1136/bmjopen-2021-049762] [Medline: 34108173]
16. Heffernan M, Andrews G, Fiatarone Singh MA, Valenzuela M, Anstey KJ, Maeder A, Maintain Your Brain Collaborative Team, et al. Maintain your brain: protocol of a 3-year randomized controlled trial of a personalized multi-modal digital health intervention to prevent cognitive decline among community dwelling 55 to 77 year olds. J Alzheimers Dis 2019 Nov 23;70(s1):S221-S237 [FREE Full text] [doi: 10.3233/JAD-180572] [Medline: 30475762]
17. Wesselman LM, Hooghiemstra AM, Schoonmade LJ, de Wit MC, van der Flier WM, Sikkes SA. Web-based multidomain lifestyle programs for brain health: comprehensive overview and meta-analysis. JMIR Ment Health 2019 Apr 09;6(4):e12104 [FREE Full text] [doi: 10.2196/12104] [Medline: 30964438]
18. Mehra S, van den Helder J, Visser B, Engelbert RH, Weijs PJ, Kröse BJ. Evaluation of a blended physical activity intervention for older adults: mixed methods study. J Med Internet Res 2020 Jul 23;22(7):e16380 [FREE Full text] [doi: 10.2196/16380] [Medline: 32459652]
19. Moorthy P, Weinert L, Schüttler C, Svensson L, Sedlmayr B, Müller J, et al. Attributes, methods, and frameworks used to evaluate wearables and their companion mHealth apps: scoping review. JMIR Mhealth Uhealth 2024 Apr 05;12:e52179 [FREE Full text] [doi: 10.2196/52179] [Medline: 38578671]
20. Arigo D, Jake-Schoffman DE, Wolin K, Beckjord E, Hekler EB, Pagoto SL. The history and future of digital health in the field of behavioral medicine. J Behav Med 2019 Feb;42(1):67-83 [FREE Full text] [doi: 10.1007/s10865-018-9966-z] [Medline: 30825090]
21. Fox S, Brown LJ, Antrobus S, Brough D, Drake RJ, Jury F, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. JMIR Mhealth Uhealth 2022 Jan 14;10(1):e24483 [FREE Full text] [doi: 10.2196/24483] [Medline: 35029539]
22. Harrington CN, Wilcox L, Connelly K, Rogers W, Sanford J. Designing health and fitness apps with older adults: examining the value of experience-based co-design. In: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2018 Presented at: PervasiveHealth '18; May 21-24, 2018; New York, NY p. 15-24 URL: https://dl.acm.org/doi/10.1145/3240925.3240929 [doi: 10.1145/3240925.3240929]
23. Lim MH, Gleeson JF, Rodebaugh TL, Eres R, Long KM, Casey K, et al. A pilot digital intervention targeting loneliness in young people with psychosis. Soc Psychiatry Psychiatr Epidemiol 2020 Jul 14;55(7):877-889. [doi: 10.1007/s00127-019-01681-2] [Medline: 30874828]
24. Audibert CE, de Moura Fereli Reis A, Zazula R, Machado RC, Guariente SM, Nunes SO. Development of digital intervention through a mobile phone application as an adjunctive treatment for bipolar disorder: MyBee project. Clin eHealth 2022 Dec;5:72-78. [doi: 10.1016/j.ceh.2022.09.001]
25. Peake E, Miller I, Flannery J, Chen L, Lake J, Padmanabhan A. Preliminary efficacy of a digital intervention for adolescent depression: randomized controlled trial. J Med Internet Res 2024 Feb 07;26:e48467 [FREE Full text] [doi: 10.2196/48467] [Medline: 38324367]
26. Ponzo S, Morelli D, Kawadler JM, Hemmings NR, Bird G, Plans D. Efficacy of the digital therapeutic mobile app BioBase to reduce stress and improve mental well-being among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17767 [FREE Full text] [doi: 10.2196/17767] [Medline: 31926063]
27. Gräfe V, Moritz S, Greiner W. Correction to: health economic evaluation of an internet intervention for depression (deprexis), a randomized controlled trial. Health Econ Rev 2020 Jul 30;10(1):24 [FREE Full text] [doi: 10.1186/s13561-020-00280-1] [Medline: 32734480]
28. Atik E, Stricker J, Schückes M, Pittig A. Efficacy of a brief blended cognitive behavioral therapy program for the treatment of depression and anxiety in university students: uncontrolled intervention study. JMIR Ment Health 2023 Aug 25;10:e44742 [FREE Full text] [doi: 10.2196/44742] [Medline: 37624631]
29. Maramis C, Moulos I, Ioakimidis I, Papapanagiotou V, Langlet B, Lekka I, et al. A smartphone application for semi-controlled collection of objective eating behavior data from multiple subjects. Comput Methods Programs Biomed 2020 Oct;194:105485. [doi: 10.1016/j.cmpb.2020.105485] [Medline: 32464588]
30. Pihlajamäki J, Männikkö R, Tilles-Tirkkonen T, Karhunen L, Kolehmainen M, Schwab U, StopDia study group. Digitally supported program for type 2 diabetes risk identification and risk reduction in real-world setting: protocol for the StopDia model and randomized controlled trial. BMC Public Health 2019 Mar 01;19(1):255 [FREE Full text] [doi: 10.1186/s12889-019-6574-y] [Medline: 30823909]
31. Rabbi M, Aung MS, Gay G, Reid MC, Choudhury T. Feasibility and acceptability of mobile phone-based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults. J Med Internet Res 2018 Oct 26;20(10):e10147 [FREE Full text] [doi: 10.2196/10147] [Medline: 30368433]

32. Schuffelen J, Maurer LF, Lorenz N, Rötger A, Pietrowsky R, Gieselmann A. The clinical effects of digital cognitive behavioral therapy for insomnia in a heterogenous study sample: results from a randomized controlled trial. Sleep 2023 Nov 08;46(11):zsad184 [FREE Full text] [doi: 10.1093/sleep/zsad184] [Medline: 37428712]
33. Upton E, Doogan C, Fleming V, Leyton PQ, Barbera D, Zeidman P, et al. Efficacy of a gamified digital therapy for speech production in people with chronic aphasia (iTalkBetter): behavioural and imaging outcomes of a phase II item-randomised clinical trial. EClinicalMedicine 2024 Apr;70:102483 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102483] [Medline: 38685927]
34. Harjumaa M, Absetz P, Ermes M, Mattila E, Männikkö R, Tilles-Tirkkonen T, et al. Internet-based lifestyle intervention to prevent type 2 diabetes through healthy habits: design and 6-month usage results of randomized controlled trial. JMIR Diabetes 2020 Aug 11;5(3):e15219 [FREE Full text] [doi: 10.2196/15219] [Medline: 32779571]
35. Chong SO, Pedron S, Abdelmalak N, Laxy M, Stephan A. An umbrella review of effectiveness and efficacy trials for app-based health interventions. NPJ Digit Med 2023 Dec 16;6(1):233 [FREE Full text] [doi: 10.1038/s41746-023-00981-x] [Medline: 38104213]
36. Barlow C. Oncology research: clinical trial management systems, electronic medical record, and artificial intelligence. Semin Oncol Nurs 2020 Apr;36(2):151005. [doi: 10.1016/j.soncn.2020.151005] [Medline: 32253050]
37. Nourani A, Ayatollahi H, Dodaran MS. A review of clinical data management systems used in clinical trials. Rev Recent Clin Trials 2019;14(1):10-23. [doi: 10.2174/1574887113666180924165230] [Medline: 30251611]
38. Clinical Trial Management System (CTMS). Definitive Healthcare Limited. URL: https://www.definitivehc.com/resources/glossary/clinical-trial-management-system [accessed 2025-06-11]
39. Drew CJ, Poile V, Trubey R, Watson G, Kelson M, Townson J, et al. Integrating technology into complex intervention trial processes: a case study. Trials 2016 Nov 17;17(1):551 [FREE Full text] [doi: 10.1186/s13063-016-1674-9] [Medline: 27855710]
40. Nourani A, Ayatollahi H, Dodaran MS. Clinical trial data management software: a review of the technical features. Rev Recent Clin Trials 2019;14(3):160-172. [doi: 10.2174/1574887114666190207151500] [Medline: 30734683]
41. Ghazali M, Ariffin NA, Omar R. User centered design practices in healthcare: a systematic review. In: Proceedings of the 3rd International Conference on User Science and Engineering. 2014 Presented at: i-USEr '15; September 2-5, 2014; Shah Alam, Malaysia p. 91-96 URL: https://ieeexplore.ieee.org/document/7002683 [doi: 10.1109/iuser.2014.7002683]
42. ISO 9241-210:2019: ergonomics of human-system interaction: part 210: human-centred design for interactive systems. International Organization for Standardization. URL: https://www.iso.org/standard/77520.html [accessed 2025-05-29]
43. Matthews T, Judge T, Whittaker S. How do designers and user experience professionals actually perceive and use personas? In: Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 1219-1208 URL: https://dl.acm.org/doi/10.1145/2207676.2208573 [doi: 10.1145/2207676.2208573]
44. Bruinsma J, Heger I, Loukas VS, Kassiotis T, Karanasiou G, Fotiadis DI, et al. Public perspectives on lifestyle-related behavior change for dementia risk reduction: an exploratory qualitative study in the Netherlands. J Alzheimers Dis 2023;95(4):1635-1642 [FREE Full text] [doi: 10.3233/JAD-230217] [Medline: 37718799]
45. Bruinsma J, Loukas VS, Kassiotis T, Heger I, Rosenberg A, Visser LN, et al. Socio-cognitive determinants of lifestyle behavior in the context of dementia risk reduction: a population-based study in the Netherlands. J Alzheimers Dis 2024;99(3):941-952 [FREE Full text] [doi: 10.3233/JAD-231369] [Medline: 38759007]
46. Peters GJ, Crutzen R. Pragmatic nihilism: how a theory of nothing can help health psychology progress. Health Psychol Rev 2017 Jun 10;11(2):103-121 [FREE Full text] [doi: 10.1080/17437199.2017.1284015] [Medline: 28110627]
47. Bartholomew Eldredge LK, Markham CM, Ruiter RA, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs: An Intervention Mapping Approach. 4th edition. New York, NY: Jossey-Bass Inc; 2016.
48. Kok G, Gottlieb NH, Peters GJ, Mullen PD, Parcel GS, Ruiter RA, et al. A taxonomy of behaviour change methods: an intervention mapping approach. Health Psychol Rev 2016 Sep;10(3):297-312 [FREE Full text] [doi: 10.1080/17437199.2015.1077155] [Medline: 26262912]
49. Interactive presentation software. Mentimeter. URL: https://www.mentimeter.com/ [accessed 2025-05-29]
50. Padlet. URL: https://de.padlet.com/ [accessed 2025-05-29]
51. Li SH, Achilles MR, Spanos S, Habak S, Werner-Seidler A, O'Dea B. A cognitive behavioural therapy smartphone app for adolescent depression and anxiety: co-design of ClearlyMe. Cogn Behav Ther 2022 Mar 16;15:e13. [doi: 10.1017/S1754470X22000095]
52. The collaborative interface design tool. Figma GmbH. URL: https://www.figma.com/ [accessed 2025-05-29]
53. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
54. Work better together with Mural's visual work platform. Mural. URL: https://mural.co/ [accessed 2023-11-06]
55. Christen OM, Mösching Y, Müller P, Denecke K, Nüssli S. Dashboard visualization of information for emergency medical services. Stud Health Technol Inform 2020 Nov 23;275:27-31. [doi: 10.3233/SHTI200688] [Medline: 33227734]

56. Bach K, Marling C, Mork PJ, Aamodt A, Mair FS, Nicholl BI. Design of a clinician dashboard to facilitate co-decision making in the management of non-specific low back pain. J Intell Inf Syst 2018 Dec 3;52(2):269-284. [doi: 10.1007/s10844-018-0539-y]
57. Hensley NB, Grant MC, Cho BC, Suffredini G, Abernathy JA. How do we use dashboards to enhance quality in cardiac anesthesia? J Cardiothorac Vasc Anesth 2021 Oct;35(10):2969-2976. [doi: 10.1053/j.jvca.2021.04.034] [Medline: 34059439]
58. R Core Team. R: a language and environment for statistical computing. The R Foundation. URL: https://www.R-project.org/ [accessed 2025-05-29]
59. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. Boca Raton, FL: CRC Press; 1996:189-194.
60. Holden RJ. A simplified System Usability Scale (SUS) for cognitively impaired and older adults. Proc Int Symp Hum Factors Ergon Healthc 2020 Sep 16;9(1):180-182. [doi: 10.1177/2327857920091021]
61. Taherdoost H. Designing a questionnaire for a research paper: a comprehensive guide to design and develop an effective questionnaire. Asian J Manag Sci 2022 Apr 25;11(1):8-16. [doi: 10.51983/ajms-2022.11.1.3087]
62. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. Waltham, MA: Elsevier Inc; 2012.
63. Welcome to FHIR®. HL7 FHIR. URL: https://www.hl7.org/fhir/? [accessed 2024-01-03]
64. Bjerke MB, Renger R. Being smart about writing SMART objectives. Eval Program Plann 2017 Apr;61:125-127. [doi: 10.1016/j.evalprogplan.2016.12.009] [Medline: 28056403]
65. Fogg BJ, Hreha J. Behavior wizard: a method for matching target behaviors with solutions. In: Proceedings of the 5th International Conference on Persuasive Technology. 2010 Presented at: PERSUASIVE '10; June 7-10, 2010; Copenhagen, Denmark p. 117-131 URL: https://link.springer.com/chapter/10.1007/978-3-642-13226-1_13 [doi: 10.1007/978-3-642-13226-1_13]
66. Wood W, Neal DT. Healthy through habit: interventions for initiating and maintaining health behavior change. Behav Sci Policy 2016 Apr 01;2(1):71-83. [doi: 10.1177/237946151600200109]
67. Stop DC. Pienet teot -sovellus. University of Eastern Finland. URL: https://sites.uef.fi/stopdia/t2d-data-pienet-teot/ [accessed 2025-06-11]
68. Kivipelto M, Solomon A, Ahtiluoto S, Ngandu T, Lehtisalo J, Antikainen R, et al. The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress. Alzheimers Dement 2013 Nov;9(6):657-665. [doi: 10.1016/j.jalz.2012.09.012] [Medline: 23332672]
69. Hilberger H, Bödenler M, Hanke S. Tracking usage patterns of a mobile application for dementia risk assessment in elderly individuals. Stud Health Technol Inform 2024 Aug 22;316:478-482. [doi: 10.3233/SHTI240452] [Medline: 39176781]
70. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
71. Hughes CP, Berg L, Danziger W, Coben LA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry 1982 Jun 29;140(6):566-572. [doi: 10.1192/bjp.140.6.566] [Medline: 7104545]
72. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
73. Mclellan S, Muddimer A, Peres SC. The effect of experience on system usability scale ratings. J Usability Stud 2012;7(2):56-67 [FREE Full text]
74. Győrffy Z, Boros J, Döbrössy B, Girasek E. Older adults in the digital health era: insights on the digital health related knowledge, habits and attitudes of the 65 year and older population. BMC Geriatr 2023 Nov 27;23(1):779 [FREE Full text] [doi: 10.1186/s12877-023-04437-5] [Medline: 38012565]
75. Ruissen MM, Torres-Peña JD, Uitbeijerse BS, Arenas de Larriva AP, Huisman SD, Namli T, POWER2DM Consortium. Clinical impact of an integrated e-health system for diabetes self-management support and shared decision making (POWER2DM): a randomised controlled trial. Diabetologia 2023 Dec 29;66(12):2213-2225 [FREE Full text] [doi: 10.1007/s00125-023-06006-2] [Medline: 37775611]
76. Wegener EK, Bergschöld JM, Whitmore C, Winters M, Kayser L. Involving older people with frailty or impairment in the design process of digital health technologies to enable aging in place: scoping review. JMIR Hum Factors 2023 Jan 27;10:e37785 [FREE Full text] [doi: 10.2196/37785] [Medline: 36705959]
77. LETHE (λήθη) – a personalized prediction and intervention model for early detection and reduction of risk factors causing dementia, based on AI and distributed Machine Learning. European Commission. URL: https://cordis.europa.eu/project/id/101017405 [accessed 2025-05-29]

Abbreviations

AI: artificial intelligence
CI/CD: Continuous Integration/Continuous Delivery or Deployment
CRF: case report form
CTMS: clinical trial management system

eCRF: electronic case report form
FHIR: Fast Healthcare Interoperability Resources
FINGER: Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability
ICT: information and communications technology
mHealth: mobile health
POWER2DM: predictive model-based decision support for diabetes patient empowerment
PRODEMOS: prevention of dementia using mobile phone applications
RCT: randomized controlled trial
SUS: System Usability Scale
UCD: user-centered design

Edited by P Jansons; submitted 19.09.24; peer-reviewed by K Kawaguchi, R Pryss; comments to author 30.03.25; revised version received 06.05.25; accepted 26.05.25; published 02.07.25.

Please cite as:
Hilberger H, Buchgraber-Schnalzer B, Huber S, Weitlaner T, Bödenler M, Abaci A, Bruinsma J, Diaz A, Guazzarini AG, Lehtisalo J, Lee S, Loukas V, Mangialasche F, Mecocci P, Ngandu T, Rosenberg A, Stögmann E, Valkonen K, Uhlik E, Untersteiner H, Kneß L, Ahammer H, Hanke S
Design of a Mobile App and a Clinical Trial Management System for Cognitive Health and Dementia Risk Reduction: User-Centered Design Approach
JMIR Aging 2025;8:e66660
URL: https://aging.jmir.org/2025/1/e66660
doi:10.2196/66660
PMID:

Network Analysis of Key Instrumental Activities of Daily Living and Cognitive Domains for Targeted Intervention in US Older Adults Without Dementia: Cross-Sectional Study

Jiaying Li[1,2], PhD; Rendong He[1,3], MSN; Erh-Chi Hsu[1], MPH; Junxin Li[1], PhD

[1]School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

[2]School of Nursing, Li Ka Shing Faculty of Medicine, University of Hong Kong, China (Hong Kong)

[3]School of Nursing, Jilin University, Changchun, China

Corresponding Author:
Junxin Li, PhD
School of Nursing, Johns Hopkins University, 525 N Wolfe St, Baltimore, MD, United States

Abstract

Background: Cognitive impairment in older adults reduces independence and raises health care costs but can be mitigated through stimulating activities. Based on network theory, intricate relationships within and between clusters of instrumental activities of daily living (IADLs) and cognitive domains suggest the existence of central IADLs and cognitive domains, as well as bridge IADLs. Modifying these can significantly enhance daily living activities and cognitive functions holistically.

Objective: This study aims to identify central IADLs (key activities within the IADL network), central cognitive domains (key domains within the cognitive network), and bridge IADLs (linking IADL and cognitive networks). These insights will inform targeted interventions to effectively improve IADL and cognitive well-being in older adults.

Methods: A cross-sectional analysis of adults aged 65 years and older in the United States focused on 5 IADLs and 6 cognitive domains from the National Health and Aging Trends Study (NHATS). Network analysis identified central and bridge variables. Nonparametric and case-dropping bootstrap methods checked network stability. Network comparison tests assessed sex differences with Benjamini-Hochberg adjustments.

Results: Of the 2239 participants, 56.4% were female (n=976). We computed and tested 3 networks: IADL, cognition, and bridge-with correlation stability coefficients of 0.67, 0.75, and 0.44, respectively (all>0.25). Meal preparation was identified as the central IADL, with a centrality index of 3.87, which was significantly higher than that of other IADLs (all P<.05). Visual attention emerged as the central cognition domain, with a centrality index of 0.86, which was significantly higher than that of other cognition domains (all P<.05). Shopping was determined to be the bridge IADL, with a centrality index of 0.41, which was significantly higher than that of other IADLs (all P<.05). Notably, gender differences emerged in the IADL network, with stronger associations between laundry and meal preparation in females (1.69 vs males: 0.74; P=.001) and higher centrality in meal preparation among females (difference=1.99; P=.007).

Conclusions: While broad enhancements in all IADL and cognitive domains are beneficial, targeting meal preparation, visual attention, and shopping may leverage their within-network influence to yield a more pronounced improvement in holistic IADL, holistic cognition, and holistic cognition function through IADL interventions among older adults. Notably, meal preparation interventions may be less effective in males, requiring tailored approaches.

(JMIR Aging 2025;8:e67632) doi:10.2196/67632

KEYWORDS

cognition function; older adults; intervention targets; elder; elderly; cognitive impairment; stimulating activity; instrumental activities of daily living; IADL; daily living activity; cognitive domain; non-demented; cognitive network; holistic cognition; holistic cognition function; network comparison; central variables; bridge variables; network analysis

Introduction

Cognitive function encompasses mental processes such as acquiring knowledge, manipulating information, and reasoning, including perception, memory, learning, attention, decision-making, and language abilities [1]. Globally, cognitive impairment in older adults has a prevalence ranging from 5.1% to 41%, with incidence rates around 53.97 per 1000 person-years [2]. This impairment not only predicts older adults' future incidence of dementia, but also significantly reduces functional independence and quality of life [2]. Beyond the individual level, economically, cognitive impairment incurs 44% higher direct medical costs and significantly increases the need for

informal care [3]. Given these consequences, early detection and proactive management are essential to mitigate impacts and prevent progression to more severe conditions and greater health care burdens.

Instrumental activities of daily living (IADLs) encompass complex tasks necessary for independent living, such as cooking, cleaning, transportation, laundry, and financial management [4]. These activities assessment sometime adjusted by country or age will include driving or medication management [5,6] and demand cognitive skills like planning, memory, and problem-solving, linking their performance closely to cognitive well-being [4]. In performing these tasks, individuals engage in practical cognitive training that builds cognitive reserve and supports overall brain health. Furthermore, regularly performing IADLs provides ongoing cognitive training that builds "cognitive reserve," a concept suggesting that mentally stimulating tasks strengthen neural connections, thereby enhancing brain resilience [7]. Although other pursuits, such as employment or volunteering, also require high-level cognition, IADLs foster daily independence in older adults, and their everyday, repetitive occurrence makes them an ideal approach for continuous cognitive training. However, current interventions usually target specific IADL domains, such as shopping or meal preparation, highlighting the challenges and resource demands of broad IADL interventions [8].

A significant research gap exists in identifying the most impactful IADL component that can efficiently enhance global cognitive function, particularly crucial in resource-limited settings. This need is supported by complex interactions between IADLs and cognition [9-12], correlations within individual IADL domains (eg, between laundry and meal preparation due to similar cognitive and physical demands) [13], and relationships within cognition domains (between psychomotor function and visual attention stemming from their joint role in tasks requiring quick coordination and responses) [14]. Notably, the cognitive demands of specific IADLs vary; for instance, shopping necessitates skills in navigation, selection, and financial transactions, whereas meal preparation involves planning, execution, and presentation. Such distinctions suggest the existence of certain IADLs that are more closely linked to multiple cognitive domains. Traditional analytic methods, which often isolate relationships or assume predictor independence, may overlook the nuanced, simultaneous interactions among nodes. In contrast, network analysis captures these complex dynamics by representing variables as nodes interconnected by edges [15]. This approach not only elucidates direct interactions but also reveals broader network structures, thereby identifying "central" nodes-those exerting significant within-group influence-and "bridge" nodes that connect disparate networks [15]. Specifically, within our framework, a central IADL is defined as the activity with the highest connectivity within the IADL network, while central cognition refers to the cognitive domain with the most extensive links. A bridge IADL, by linking the IADL and cognitive networks, may have an outsized impact on overall function when its performance changes. Such insights can inform target interventions and strategic resource allocation aimed at enhancing both daily living activities and cognitive function [15]. Furthermore, gender-specific differences in cognitive decline and IADL performance further complicate this landscape. Women generally demonstrate superior executive function and memory; however, their executive function appears to decline more rapidly than that of men, while memory trajectories remain similar between sexes. In contrast, difficulties in financial management and medication adherence are more predictive of dementia in men, whereas transportation challenges serve as stronger indicators in women [16,17]. Consequently, a nuanced exploration of these differences is critical for developing targeted intervention strategies.

This study aimed to identify the central IADL and cognition domains within their respective networks, pinpoint the bridge IADL most substantially linked to overall cognitive function, and examine sex differences in these variables. We hypothesized the existence of central, bridge variables and sex-based differences affecting them. The findings are expected to reveal network dynamics, pinpoint key intervention targets for effectively enhancing holistic IADL and cognitive functions in the elderly, and indicate the necessity of sex-specific interventions.

Methods

Study Design and Data Source

This cross-sectional analysis used waves 11 and 12 (2021 - 2022) of the National Health and Aging Trends Study (NHATS), a nationally representative longitudinal database of Medicare beneficiaries aged 65 years and older. Waves 11 and 12 of NHATS were chosen for their comprehensive 6-domain cognitive assessment, unlike earlier waves that measured only episodic memory, executive function, and orientation. Data were gathered during in-home interviews by NHATS interviewers. For those included in both waves, only data from wave 12 were retained to ensure the latest cognitive assessments were used. The reporting of this study followed the CHERRIES checklist [18].

Participants and Sample Size Calculation

Eligible participants were cognitively intact individuals aged 65 years and older, not residing in nursing homes, and without signs of cognitive impairment. Cognitive intactness was determined by absence of potential or probable dementia. According to previous NHATS literature, potential dementia was defined by scores≤1.5SDs below the mean in one cognitive domain, while probable dementia was indicated by similar scores in at least two domains, meeting AD8 criteria, or having a clinical dementia diagnosis [19]. To ensure adequate statistical power, our network analysis of 11 nodes and 55 edges required a minimum of 165 participants, adhering to the 3-participants-per-parameter rule [20].

Measures

Cognitive performance was assessed across six domains: (1) Episodic memory, scored 0 - 20 from immediate and delayed recall of 10 words; (2) executive function, scored 0 - 5 by the clock drawing test; (3) orientation, scored 0 - 8 from knowledge of the current date and names of the president or vice president; (4) psychomotor function, measured by reaction speed in log-transformed milliseconds from the cogstate detection task,

where participants respond when a card turns face up; (5) visual attention, assessed by reaction speed for correct responses in log-transformed milliseconds from the cogstate identification task, where participants decide if a card is red or black; and (6) working memory, evaluated by accuracy in the cogstate one card back task, asking participants to remember if they have seen a card before. Cogstate, a tablet-based test used since wave 11 in NHATS, includes card detection, identification, and one card back activities, expanding cognitive assessments beyond traditional tests. This computerized assessment has been validated against traditional paper-based cognitive tests, demonstrating adequate reliability and validity in differentiating adults with cognitive impairment from those without [21]. Further details are available in the NHATS Cogstate user guide [22].

IADLs were assessed via self-reports on managing medication, laundry, shopping, meal preparation, and banking. Participants reported performance over the past month using 5 options: “1” Did not do by self last month; “2” Did by self last month with no difficulty; “3” Did by self last month with difficulty; “4” -Don’t know or refuse, with no difficulty; and “5” Don’t know or refuse, with difficulty. We dichotomized responses rather than using a Likert scale because the options do not form a natural continuum, but instead distinguish independent performance from any difficulty. Participants with no difficulty (responses “2” or “4”) were classified as “no difficulty,” while those with difficulty or inability (responses “1,” “3,” or “5”) were classified as “difficulty.” This binary approach preserves the key distinction in functional independence and aligns with previous research [23].

Sociodemographic variables included age (70 - 74, 75 - 79, 80 - 84, 85 - 89, and 90+ years), sex (female or male), living arrangement (alone, with others), race (Hispanic, non-Hispanic Black, non-Hispanic White, and other non-Hispanic), income status (poverty, low income, and normal), and self-rated health (poor, fair, good, very good, and excellent). Income status was defined according to federal guidelines from the Office of the Assistant Secretary for Planning and Evaluation at the US Department of Health and Human Services (ASPE HHS, 2024 version) as follows:≤100% Federal Poverty Level (FPL), >100% to≤200% FPL, and>200% FPL, replacing the previous labels of “poverty,” “low income,” and “normal.” Self-rated health was measured using a single 5-point Likert-scale item (1=poor, 5=excellent): “Would you say that, in general, your health is poor, fair, good, very good, or excellent?”

Statistical Analysis

Data were organized in a Microsoft Excel database and subjected to rigorous quality control checks. The analysis was carried out using R statistical software (version 4.1.1; R Core Team). Descriptive statistics summarized participant demographics and performance in IADLs and cognitive functions. Continuous variables were checked for normality with P-P plots and described using mean and SD; categorical variables were presented as frequencies and percentages. To maintain consistency across all nodes within the IADL and cognition networks, necessary reverse coding adjustments were made to ensure that higher scores consistently indicate diminished capabilities. Network analysis proceeded through 5 phases: evaluating topological overlap, estimating the network, assessing network stability, calculating centrality and bridge centrality indices, and conducting network comparison tests.

Checking Topological Overlap

We used the R package *network tools*’ goldbricker function to identify unique variables and avoid artificial relationships from symptom similarity in our network analysis. A significance threshold of 0.25 and a *P* value<.01 determined statistical significance [24].

Network Estimation

We developed 3 networks for our study: an IADL network for all 5 IADL domains, a cognitive network for all 6 cognitive domains, and a bridge network linking both. Each network consisted of nodes (items within each domain) and edges (relationships between items). For the cognition network with continuous data, we applied the EBICglasso method, which used the Extended Bayesian Information Criterion (EBIC) with the least absolute shrinkage and selection operator (LASSO) for partial correlation analysis, reducing confounding by shrinking coefficients and zeroing smaller correlations. The IADL network, based on binary data, was analyzed using the IsingFit method, which used logistic regression to adjust node states and determine conditional probabilities. The bridge network was assessed using the “mgm” method, designed for mixed data types, using conditional independence tests tailored to heterogeneous data. Network visualization was performed using R packages *bootnet* and *qgraph*, where edge thickness represented association strength—blue for positive and red for negative associations [25].

Network Stability

The *bootnet* package was used to assess edge and centrality stability in each network [25]. Edge stability was evaluated through nonparametric bootstrap, with 95% CIs reflecting accuracy; narrower CIs indicate higher network reliability [25]. Centrality stability was measured with a case-dropping subset bootstrap, as reflected by the correlation stability coefficient (CS-C); values above 0.25, ideally over 0.5, denote optimal stability [25].

Central Node, Centrality, Bridge Node, and Bridge Centrality

A central node in a network has substantial influence due to its extensive connections with other nodes [26]. Bridge nodes connect different communities or clusters within a network, facilitating interactions that would otherwise not occur [27]. Centrality measures in network analysis typically include strength, betweenness, closeness, and expected influence; however, due to the instability of betweenness and closeness, and because strength ignores negative edges (summing only absolute values), we exclusively used expected influence for central nodes and bridge expected influence for bridge nodes [28]. The expected influence index, which accounts for both positive and negative edge values, was calculated using the *qgraph* package in R [26,27]. Similarly, the top bridge node was identified through the highest bridge expected influence (1-step) index, which sums the edge values connecting the node

to those outside its community, determined by the *networktools* package in R [26,27]. Differences in node centrality were analyzed using Wilcoxon tests with 1000 bootstrapped indices from the bootnet package in R, applying Benjamini-Hochberg corrections for multiple comparisons [29].

Network Comparison Test

To analyze gender differences across 3 networks, we used the *network comparison test* package in R. This involved performing both a network invariance test, which identified significant variations in edges among subgroup networks, and a global strength invariance test, which evaluated the total weighted sum of all edges to measure the connection strength among network variables. Should the network invariance test yield significant results, we then conducted an edge invariance test to pinpoint specific edge pairs that varied between subgroups. In addition, we compared node centrality between men and women. To adjust for multiple comparisons at the level of individual edges and centralities, the Benjamini-Hochberg correction method was used.

Ethical Considerations

This secondary analysis of the NHATS dataset relies on publicly available data. The original data collection, which obtained informed consent from all participants, was approved by the Johns Hopkins University IRB. As no restricted data were used, further IRB review was not required.

Results

Sample Characteristics and Descriptions of IADL and Cognitive Domains

Of the 2239 participants (1194 from wave 12 and 245 from wave 11), 1263 were female (56.41%), and 1720 (76.82%) were White. The predominant age group was 75‐79 years, representing 748 participants (33.41%). Detailed sociodemographic data, as well as descriptions and abbreviations of the IADL items and cognitive domains, are provided in Table 1.

Table . Demographics, descriptions, and abbreviations of instrumental activities of daily living (IADL) and cognitive domain items (N=2239).

Variables	Results	Range
Demographics, n (%)		
Sex		
Female	1263 (56.41)	—[a]
Male	976 (43.59)	—
Age group		
70 - 74 years	214 (9.56)	—
75 - 79 years	748 (33.41)	—
80 - 84 years	620 (27.69)	—
85 - 89 years	395 (17.64)	—
90+ years	262 (11.70)	—
Self-rated health		
Poor	69 (3.08)	—
Fair	431 (19.25)	—
Good	893 (39.88)	—
Very good	676 (30.19)	—
Excellent	169 (7.55)	—
Missing value	1 (0.04)	—
Race		
Non-Hispanic others	41 (1.83)	—
Non-Hispanic Black	375 (16.75)	—
Hispanic	81 (3.62)	—
Non-Hispanic White	1720 (76.82)	—
Missing	22 (0.98)	—
Living arrangement		
Alone	809 (36.13)	—
Living with someone	1430 (63.87)	—
Income status		
Poverty	859 (38.37)	—
Low income	413 (18.45)	—
Normal	967 (43.19)	—
Description of items (abbreviations in networks), n (%)		
Difficulty in managing medication (I1)		
No	1798 (80.30)	—
Yes	441 (19.70)	—
Difficulty in managing laundry (I2)		
No	1321 (59)	—
Yes	918 (41)	—
Difficulty in managing shopping (I3)		
No	1214 (54.22)	—
Yes	1025 (45.78)	—
Difficulty in managing meal (I4)		
No	1373 (61.32)	—

Variables	Results	Range
Yes	866 (38.68)	—
Difficulty in managing banking (I5)		
No	1505 (67.22)	—
Yes	734 (32.78)	—
Episodic memory (C1), mean (SD)	10.15 (3.21)	0-9
Executive function (C2), mean (SD)	0.84 (0.89)	0-5
Orientation (C3), mean (SD)	0.96 (1.33)	0-8
Psychomotor function (C4), mean (SD)	2.66 (0.13)	2.36-3.26
Visual attention (C5), mean (SD)	2.82 (0.10)	2.57-3.41
Working memory (C6)	0.34 (0.25)	0-1.57

[a]Not applicable.

Items Remained After Redundancy Check

The Goldbricker analysis confirmed no redundancy in the IADL and cognitive domains, thus all items were retained.

Stability of IADLs, Cognition, and the Bridge Networks

The bootstrapped 95% CI analysis revealed narrow CIs across all 3 networks (IADL, cognition, and bridge), indicating precise edge-weight estimates (Figures S1, S3, and S5 in Multimedia Appendix 1). In addition, CS-C values for the IADL, cognition, and bridge networks were 0.67, 0.75, and 0.44, respectively, all surpassing the recommended threshold of 0.25, confirming the networks' interpretability and reliability (Figures S2, S4, and S6 in Multimedia Appendix 1).

IADL Network

Figure 1A illustrates the IADL network, with all edges (10/10, 100%) nonzero, indicating strong connectivity among nodes. The most robust connections were between I3 and I4 (I3: shopping-I4:meal, edge weight 1.38), I2 and I4 (I2: laundry-I4:meal, 1.33), and I1 and I5 (I1: medication- I5: banking, 1.08). Logistic regression coefficients for other edges are detailed in Table S1 in Multimedia Appendix 1. Figure 2A shows that node I4 (meal) had the highest expected influence of 3.87. Figure 2B's centrality bootstrapped difference test highlights I4's (meal's) significantly higher influence (all $P<.05$ after Benjamini-Hochberg corrections), underscoring its central role in the network.

Figure 1. Network structure of instrumental activities of daily living (IADL) network, cognition network, and the bridge network.

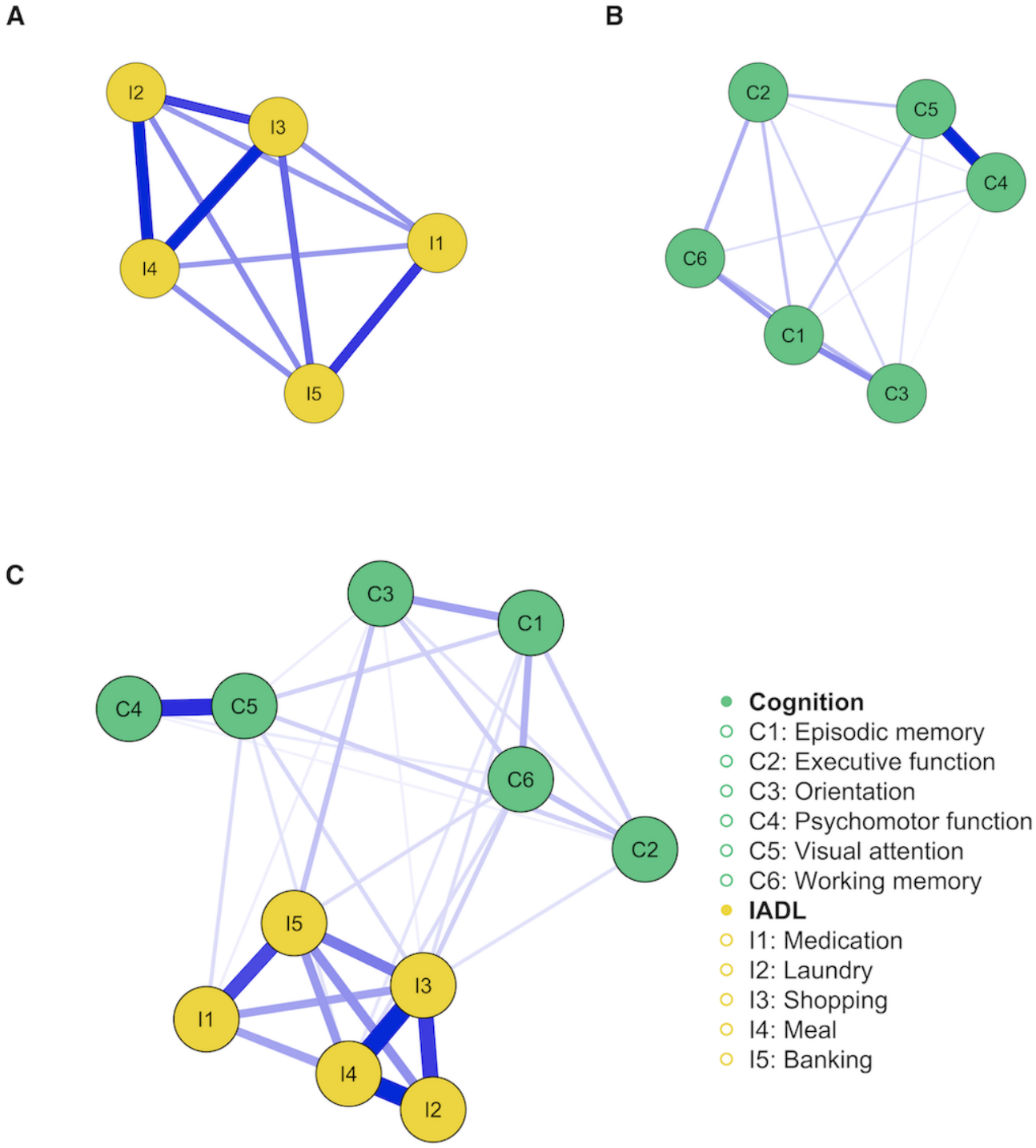
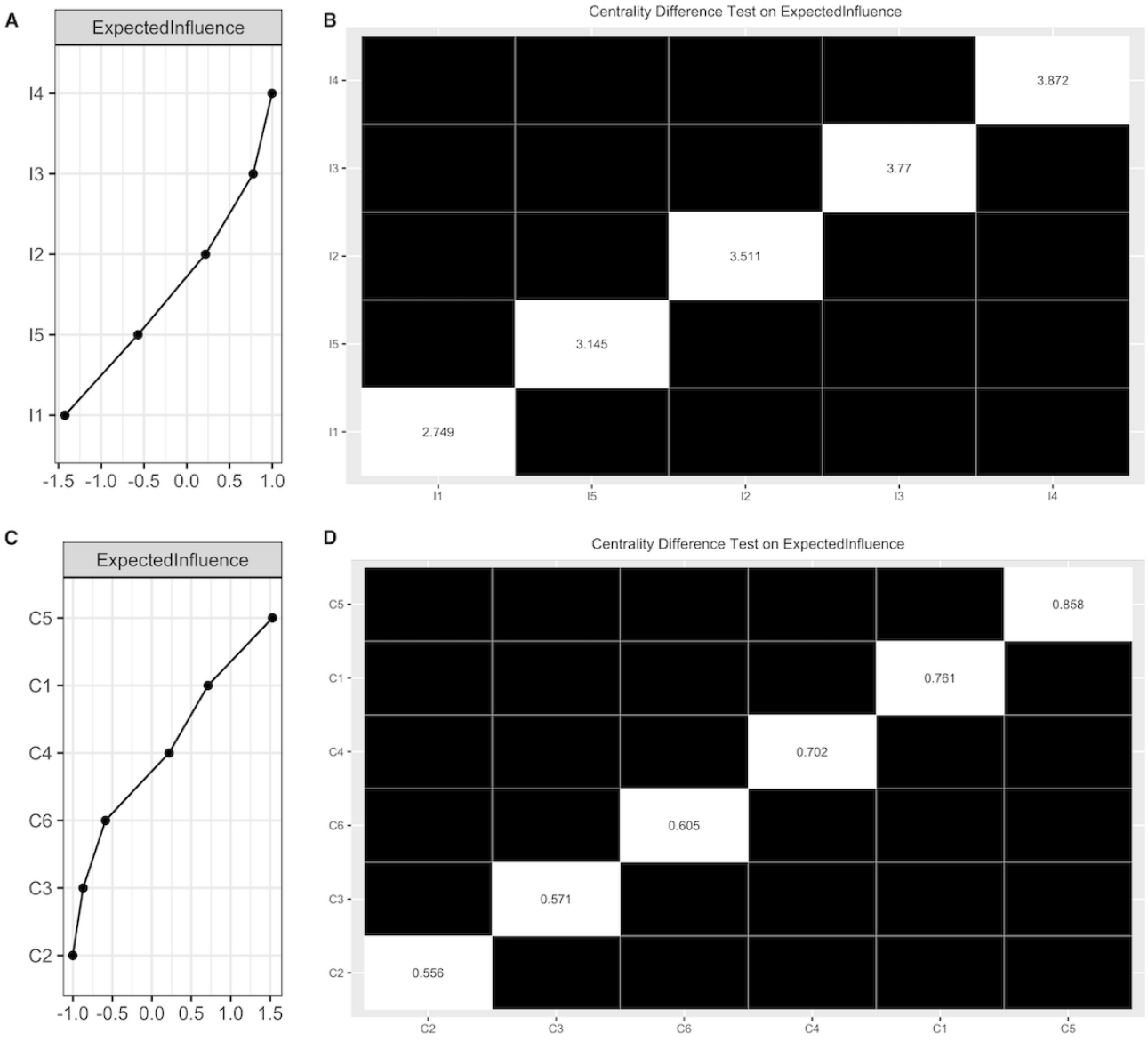


Figure 2. Expected influence centrality index and centrality bootstrapped difference tests for variables in the IADL network (Pane A and Panel B) and the cognition network (Panel C and Panel D). Gray boxes indicate variables that do not significantly differ from one-another. Black boxes represent variables that differ significantly from one another (α=.05). White boxes show the values of bridge expected influence. I1: difficulty in managing medication; I2: difficulty in managing laundry; I3: difficulty in managing shopping; I4: difficulty in managing meals; I5: difficulty in managing banking; C1: episodic memory; C2: executive function; C3: orientation; C4: psychomotor function; C5: visual attention; C6: working memory.



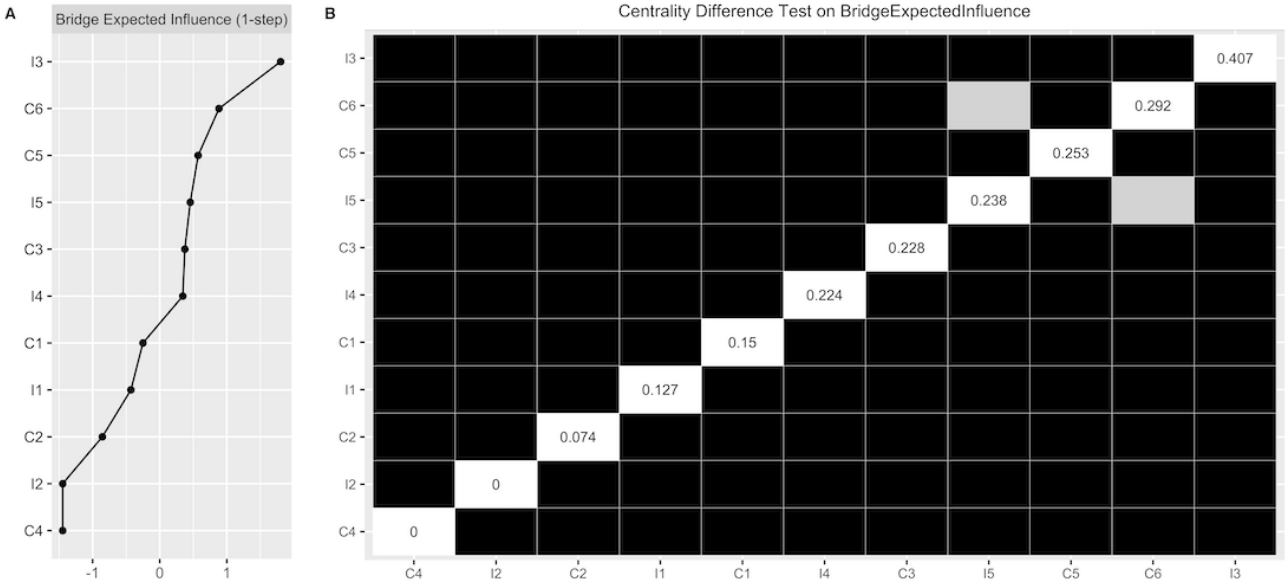
Cognition Network

Figure 1B shows the cognition network structure, with 14/15 edges (93.33%) nonzero, reflecting strong connectivity among nodes. The strongest connections were C4-C5 (C4: psychomotor function-C5: visual attention; edge weight 0.54), C1-C3 (C1: episodic memory-C3: orientation, 0.25), and C1-C6 (C1: episodic memory-C6: working memory, 0.22). The partial correlation matrix for other edges is in Table S2 in Multimedia Appendix 1. Figure 2C displays the expected influence index for all nodes, with C5 (visual attention) having the highest at 0.86. Figure 2D's centrality bootstrapped difference test underscores C5's (visual attention's) significant influence within the cognitive domains (all P<.05 after Benjamini-Hochberg corrections).

Bridge Network

Figure 1C shows the bridge network between IADL and cognition, with 34/55 edges (61.82%) nonzero, indicating strong connectivity. Details on all edges are in Table 3 in Multimedia Appendix 1. Figure 3A highlights the bridge expected influence index, with I3 (shopping) recording the highest at 0.41, followed by I5 (banking, 0.24) and I4 (meal, 0.22). Significant bridge edges include I3-C6 (I3: shopping-C6: working memory, edge weight: 0.12), I5-C3 (I5: banking-C3: orientation, edge weight: 0.15), and I4-C6 (I4: meal-C6: working memory, edge weight: 0.09). Figure 3B's centrality bootstrapped difference test confirms I3's (shopping's) prominent role in connecting IADL and cognitive domains (all P<.05 after Benjamini-Hochberg corrections).

Figure 3. Bridge expected influence centrality index (Panel A) and centrality bootstrapped difference tests (Panel B) for variables in the bridge network. Gray boxes indicate variables that do not significantly differ from one-another. Black boxes represent variables that differ significantly from one another (α=.05). White boxes show the values of bridge expected influence. I1: difficulty in managing medication; I2: difficulty in managing laundry; I3: difficulty in managing shopping; I4: difficulty in managing meals; I5: difficulty in managing banking; C1: episodic memory; C2: executive function; C3: orientation; C4: psychomotor function; C5: visual attention; C6: working memory.



Sex Differences in Networks

Network invariance and global strength tests revealed no significant sex differences in the cognition network (*M*=0.07, *P*=.64; *S*=0.10, *P*=.21). Conversely, significant sex differences were evident in the IADL and bridge networks, confirmed by network invariance tests (IADL: *M*=0.95; *P*=.005 and bridge: *M*=0.49; *P*=.002) and global strength tests (IADL: *S*=1.94; *P*=.001 and bridge: *S*=1.79; *P*=.002). Following these findings, further edge and centrality invariance tests were conducted for the IADL and bridge networks between sexes. In the IADL network, edge invariance tests identified significant sex disparities between I2 (laundry) and I4 (meal), with females demonstrating stronger associations (female 1.69, male 0.74; *P*=.001) and greater centrality in I4 (meal: difference=1.99; *P*=.007). However, in the bridge network, no significant sex differences were detected through further edge and centrality invariance tests (all *P*>.05 after Benjamini–Hochberg corrections).

Discussion

Principal Findings

Using network analysis, the main findings of this study elucidates the detailed interactions within and between IADL and cognitive domains in older adults, identifying 3 key variables as predictive markers and potential intervention targets for enhancing global IADL and cognitive function, while also noting sex differences. First, meal preparation difficulty and visual attention are central nodes within their respective IADL and cognition networks, with higher levels predictive of better functionality; targeted modifications could significantly improve overall functionality. Secondly, shopping difficulty within the IADL network has the strongest association with global cognition. Given the cause-and-effect relationship between IADL and cognitive function, early interventions targeting shopping difficulties could effectively boost global cognition. Finally, a sex difference was observed, with meal preparation exerting a greater influence in the IADL network among females than males, suggesting its higher predictive relevance and intervention efficacy for females.

Within the IADL network, the biggest 3 connections include those between shopping and meal preparation, laundry and meal preparation, and medication management and banking. The strong link between shopping and meal preparation is due to their shared planning, organization, and physical demands necessary for food tasks [30-32]. Likewise, laundry and meal preparation share organizational and physical demands [13]. The link between medication management and banking stems from their reliance on the same executive functions and working memory [33,34]. Importantly, meal preparation is central in the IADL network, perhaps due to impacts on other domains like medication, shopping, and banking owning to necessary executive skills such as multitasking and planning [35,36]. Data from our cross-sectional analysis indicate that meal preparation is the most central node within the IADL network. This finding suggests that same unite improvement in meal preparation ability can yield the largest overall enhancement in IADL performance relative to other tasks. Therefore, targeted interventions—such as the integration of assistive kitchen technologies or cognitive orthoses—may be especially effective in promoting independent cooking and, by extension, broader functional independence [37,38].

In the cognition network, the strongest connections are between psychomotor function and visual attention, and between episodic memory, orientation, and working memory. The link between psychomotor and visual attention is due to their roles in coordinated, rapid response tasks [14]. Significant correlations also exist between episodic memory and gray matter volume in the bilateral hippocampus and parahippocampal gyrus, which are key for orientation [39]. Furthermore, the association

between episodic memory and working memory is supported by evidence that working memory capacity and prefrontal cortex executive functions are essential for episodic memory formation and retrieval [40]. Importantly, visual attention, the most central node in the cognitive network, is crucial for selective focus necessary for advanced cognitive processes like reasoning and problem-solving [41], and it enhances complex task execution through interactions with working memory [42]. Studies show that enhancing visual attention can lead to sustained improvements in cognitive performance [43,44], supporting its pivotal role in overall cognition and highlighting the need for targeted interventions to enhance focus and processing of visual stimuli. Potential interventions include computer-based training programs, video games, mindfulness exercises, and virtual reality applications specifically designed for visual tracking tasks, all of which may contribute to enhanced cognitive function [45-48].

In the bridge network, which integrates both IADLs and cognition, connectivity is notably high, with 61.82% (34/55) of edges being nonzero. This finding indicates a robust interconnection among multiple cognitive domains and IADLs, suggesting that daily tasks rely on a synergy of cognitive processes rather than on any single discrete skill [49,50]. Such an observation aligns with mounting evidence that real-life cognition operates as an integrated set of processes-often termed functional cognition [51,52]. While traditional neuropsychological models emphasize isolated cognitive constructs, functional cognition highlights how domains such as attention, executive function, and memory converge to support everyday activities [53]. By adopting this framework, our findings on the specific links (“edges”) between each cognitive domain and each IADL can offer theoretical guidance for real-world functional cognition rehabilitation or training aimed at improving IADL performance. Importantly, the prominent bridge edge was identified between shopping and working memory. When considering the overall impact on global cognition, shopping ranked first as the bridge IADL, followed by banking and meal preparation. These latter activities are also highly demanding cognitively and should be targeted in interventions. However, if resources are limited, prioritizing shopping may yield the greatest benefits in enhancing overall cognitive function. Notably, shopping functions as a bridge IADL because it draws on a broad range of cognitive skills: episodic memory for recalling past purchases and layouts; executive function for planning and budgeting; spatial orientation for navigation; psychomotor skills for handling products and carts; visual attention for identifying items; and working memory for tracking purchases and costs [33,54-57]. To leverage its bridge position, shopping tasks could be integrated into routine cognitive assessments, and regular shopping activities could be encouraged to maintain cognitive function. Virtual reality simulations of shopping tasks, artificial intelligence-powered service robotics, and the use of audio recorders as assistive technology to enhance shopping independence among older adults [58,59], can leverage the bridge IADL role of shopping to effectively improve global cognitive function.

Sex differences within the IADL network show females with a stronger association between laundry and meal preparation and a higher centrality of meal preparation. These patterns likely result from societal norms assigning women more IADL responsibilities, especially laundry and meal preparation [17,60,61]. The elevated centrality of meal preparation in women’s IADL networks suggests a ripple effect where challenges in meal preparation deplete time management, mental energy, and physical resources, reducing efficiency in other tasks. This ripple effect also explains the stronger association between meal preparation and laundry among women compared with men. Previous studies indicate that IADLs do not measure equivalently for men and women [61]; our study finds that meal preparation has a lower predictive value for overall IADL function in males than in females. Interventions to improve IADL performance through meal preparation should be tailored with an awareness that these activities vary in importance and difficulty between sexes.

Limitations

Several limitations are worth mentioning. First, the assessment of IADL relied on self-reported data, which, while expedient, may compromise reliability and necessitate cautious interpretation of the findings. Future research should use more objective measures to validate these results. Second, the cross-sectional design precludes the establishment of causality and does not capture the temporal dynamics between IADL capabilities and cognitive function, underscoring the need for longitudinal approaches. Third, we excluded older adults with dementia, as dementia-related deficits may mask subtle variations in both cognition and IADL performance, which is critical for identifying central or bridge nodes in network analysis. Consequently, our findings should be interpreted primarily for community-dwelling older adults with relatively preserved cognitive function. Future research should incorporate participants with more severe cognitive impairments to validate whether our findings remain consistent. Fourth, although we set eligibility at ≥65 years, all final participants were aged 70+ years, likely because those aged 65 - 69 years in earlier waves either dropped out or turned 70 years old by Waves 11 - 12. This may limit the generalizability of our findings to younger older adults. Fifth, our study relies on the NHATS dataset, which limits the assessment to a select set of IADL and cognitive domains. Sixth, one limitation of our study is that the network analysis did not adjust for external confounders (eg, age and cultural background). Network models focus solely on the relationships among the included nodes and do not account for factors outside the network. Future research might address this limitation by residualizing each node on confounders before constructing the network, or by using subgroup or multigroup analyses to examine how these factors influence network structure. While these measures capture key aspects of functioning, we acknowledge that not including additional, more nuanced domains may affect the generalizability of our findings. Finally, the generalizability of our findings is limited to the American population studied; it remains unclear if these results can be generalized to populations with differing cultural, economic, or health system backgrounds. Further studies should

expand the demographic scope to determine if these findings hold across diverse populations.

Conclusions

The central IADL, central cognitive domain, and bridge IADL connecting global cognition were meal preparation, visual attention, and shopping, respectively, underscoring the need for targeted interventions to maximize resource efficiency and effectiveness. Specifically, enhancing meal preparation in older adults may significantly boost holistic IADL capabilities through interventions such as cooking classes, nutritional education, and tailored tools, along with support services such as interdisciplinary collaboration, caregiver training, and smart appliances. Similarly, focusing on visual attention training through methods such as computer-based programs, neurofeedback, and mindfulness exercises may substantially improve global cognitive function. Given the link between IADL performance and cognitive function, interventions centered on shopping are expected to be highly effective. This can be achieved by integrating shopping tasks into cognitive assessments and promoting regular shopping activities. Technological aids such as GPS, virtual reality simulations, and caregiver education on the cognitive benefits of shopping can further support elderly care and quality of life. In addition, observed sex differences suggest that meal preparation interventions may vary in effectiveness, with potentially lower efficacy among males, highlighting the need for tailored strategies to maximize outcomes.

Acknowledgments

This study was supported by funding from the National Institute on Aging (grant R01AG080613).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables S1-S3 and Figures S1-S6.
[DOCX File, 6100 KB - aging_v8i1e67632_app1.docx]

References

1. Kiely KM. Cognitive function. In: Michalos AC, editor. Encyclopedia of Quality of Life and Well-Being Research: Dordrecht: Springer Netherlands; 2014:974-978. [doi: 10.1007/978-94-007-0753-5_426]
2. Pais R, Ruano L, P Carvalho O, Barros H. Global cognitive impairment prevalence and incidence in community dwelling older adults—a systematic review. Geriatrics (Basel) 2020 Oct 27;5(4):84. [doi: 10.3390/geriatrics5040084] [Medline: 33121002]
3. Zhu CW, Sano M, Ferris SH, Whitehouse PJ, Patterson MB, Aisen PS. Health-related resource use and costs in elderly adults with and without mild cognitive impairment. J Am Geriatr Soc 2013 Mar;61(3):396-402. [doi: 10.1111/jgs.12132] [Medline: 23414481]
4. Guo HJ, Sapra A. Instrumental Activity of Daily Living: StatPearls Publishing; 2023.
5. Toups R, Chirles TJ, Ehsani JP, et al. Driving performance in older adults: current measures, findings, and implications for roadway safety. Innov Aging 2022;6(1):igab051. [doi: 10.1093/geroni/igab051] [Medline: 35028434]
6. Sino CGM, Sietzema M, Egberts TCG, Schuurmans MJ. Medication management capacity in relation to cognition and self-management skills in older people on polypharmacy. J Nutr Health Aging 2014 Jan;18(1):44-49. [doi: 10.1007/s12603-013-0359-2] [Medline: 24402388]
7. Li X, Song R, Qi X, et al. Influence of cognitive reserve on cognitive trajectories. Neurology (ECronicon) 2021 Oct 26;97(17):e1695-e1706. [doi: 10.1212/WNL.0000000000012728] [Medline: 34493618]
8. Buele J, Varela-Aldás JL, Palacios-Navarro G. Virtual reality applications based on instrumental activities of daily living (iADLs) for cognitive intervention in older adults: a systematic review. J Neuroeng Rehabil 2023 Dec 19;20(1):168. [doi: 10.1186/s12984-023-01292-8] [Medline: 38110970]
9. Huang S, Zhong W, Cheng Q, Shuai Y, Zhu J, Diao J. Instrumental activities of daily living function and cognitive status among Chinese older adults: a serial multiple mediation model. Front Public Health 2024;12:1378979. [doi: 10.3389/fpubh.2024.1378979] [Medline: 38756886]
10. Toth C, Tulliani N, Bissett M, Liu KP. The relationship between cognitive function and performance in instrumental activities of daily living in older adults. British Journal of Occupational Therapy 2022 Feb;85(2):120-129. [doi: 10.1177/03080226211008722]
11. Romero-Ayuso D, Castillero-Perea Á, González P, et al. Assessment of cognitive instrumental activities of daily living: a systematic review. Disabil Rehabil 2021 May;43(10):1342-1358. [doi: 10.1080/09638288.2019.1665720] [Medline: 31549907]
12. Lau SCL, Connor LT, Heinemann AW, Baum CM. Cognition and daily life activities in stroke: a network analysis. OTJR (Thorofare N J) 2022 Oct;42(4):260-268. [doi: 10.1177/15394492221111730] [Medline: 35838366]

13. Jefferson AL, Paul RH, Ozonoff A, Cohen RA. Evaluating elements of executive functioning as predictors of instrumental activities of daily living (IADLs). Arch Clin Neuropsychol 2006 May;21(4):311-320. [doi: 10.1016/j.acn.2006.03.007] [Medline: 16814980]
14. Hemager N, Vangkilde S, Thorup A, et al. Visual attention in 7-year-old children at familial high risk of schizophrenia or bipolar disorder: The Danish high risk and resilience study VIA 7. J Affect Disord 2019 Nov 1;258:56-65. [doi: 10.1016/j.jad.2019.07.079] [Medline: 31394459]
15. Hevey D. Network analysis: a brief overview and tutorial. Health Psychol Behav Med 2018 Sep 25;6(1):301-328. [doi: 10.1080/21642850.2018.1521283] [Medline: 34040834]
16. Levine DA, Gross AL, Briceño EM, et al. Sex differences in cognitive decline among US adults. JAMA Netw Open 2021 Feb 1;4(2):e210169. [doi: 10.1001/jamanetworkopen.2021.0169] [Medline: 33630089]
17. Ryu HJ, Moon Y. Gender differences in items of the instrumental activities of daily living in mild cognitive impairment and alzheimer's disease dementia. Dement Neurocogn Disord 2024 Apr;23(2):107-114. [doi: 10.12779/dnd.2024.23.2.107] [Medline: 38720826]
18. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
19. Kasper JD, Freedman VA, Spillman BC. Classification of persons by dementia status in the national health and aging trends study. Technical paper.: National Health and Aging Trends Study (NHATS); 2013.
20. Li J, Fong DYT, Lok KYW, et al. Key lifestyles and interim health outcomes for effective interventions in general populations: A network analysis of A large international observational study. J Glob Health 2023 Oct 20;13:04125. [doi: 10.7189/jogh.13.04125] [Medline: 37861130]
21. Lim YY, Ellis KA, Harrington K, et al. Use of the CogState Brief Battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. J Clin Exp Neuropsychol 2012;34(4):345-358. [doi: 10.1080/13803395.2011.643227] [Medline: 22248010]
22. Hu M, Freedman VA. National hHealth and Aging Trends Study (NHATS) cogstate user guide. : Johns Hopkins University School of Public Health; 2022 URL: https://nhats.org/sites/default/files/2022-07/NHATS%20Cogstate%20User%20Guide.pdf [accessed 2025-03-11]
23. Edwards RD, Brenowitz WD, Portacolone E, et al. Difficulty and help with activities of daily living among older adults living alone with cognitive impairment. Alzheimers Dement 2020 Aug;16(8):1125-1133. [doi: 10.1002/alz.12102] [Medline: 32588985]
24. Hittner JB, May K, Silver NC. A Monte Carlo evaluation of tests for comparing dependent correlations. J Gen Psychol 2003 Apr;130(2):149-168. [doi: 10.1080/00221300309601282] [Medline: 12773018]
25. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. Behav Res Methods 2018 Feb;50(1):195-212. [doi: 10.3758/s13428-017-0862-1] [Medline: 28342071]
26. Fonseca-Pedrero E. Network analysis in psychology. Papeles del Psicólogo 2018;39(1):1-12. [doi: 10.23923/pap.psicol2018.2852]
27. Jones PJ, Ma R, McNally RJ. Bridge centrality: a network approach to understanding comorbidity. Multivariate Behav Res 2021;56(2):353-367. [doi: 10.1080/00273171.2019.1614898] [Medline: 31179765]
28. Li J, Fong DYT, Lok KYW, et al. Country-specific key lifestyle factors and health outcomes for resource allocation in the general population: a network analysis across 29 countries. J Glob Health 2025 Jan 10;15:04011. [doi: 10.7189/jogh.15.04011] [Medline: 39791329]
29. Li J, Fong DYT, Lok KYW, et al. Key lifestyles and health outcomes across 16 prevalent chronic diseases: A network analysis of an international observational study. J Glob Health 2024;14. [doi: 10.7189/jogh-14-04068]
30. Lawton M, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Nurs Res 1970;19(3):278. [doi: 10.1097/00006199-197005000-00029]
31. Giere S, Rudrud E, McKay S. A functional approach to meal planning, preparation, and shopping skills. Australia and New Zealand Journal of Developmental Disabilities 1989 Jan;15(2):81-97. [doi: 10.1080/07263868900033551]
32. Davies G. Time, food shopping and food preparation: some attitudinal linkages. British Food Journal 1997 Apr 1;99(3):80-88. [doi: 10.1108/00070709710168914]
33. Engle RW. Working memory capacity as executive attention. Curr Dir Psychol Sci 2002 Feb;11(1):19-23. [doi: 10.1111/1467-8721.00160]
34. Carpenter PA, Just MA, Reichle ED. Working memory and executive function: evidence from neuroimaging. Curr Opin Neurobiol 2000 Apr;10(2):195-199. [doi: 10.1016/s0959-4388(00)00074-x] [Medline: 10753796]
35. Craik FIM, Bialystok E. Planning and task management in older adults: Cooking breakfast. Mem Cognit 2006 Sep;34(6):1236-1249. [doi: 10.3758/BF03193268]
36. Doherty TA, Barker LA, Denniss R, Jalil A, Beer MD. The cooking task: making a meal of executive functions. Front Behav Neurosci 2015;9:22. [doi: 10.3389/fnbeh.2015.00022] [Medline: 25717294]
37. Kuoppamäki S, Tuncer S, Eriksson S, McMillan D. Designing kitchen technologies for ageing in place: a video study of older adults' cooking at home. Proc ACM Interact Mob Wearable Ubiquitous Technol 2021;5(2):1-19. [doi: 10.1145/3463516]

38. Yaddaden A, Couture M, Gagnon-Roy M, et al. Using a cognitive orthosis to support older adults during meal preparation: Clinicians' perspective on COOK technology. J Rehabil Assist Technol Eng 2020;7:2055668320909074. [doi: 10.1177/2055668320909074] [Medline: 32435504]
39. Maillet D, Rajah MN. Association between prefrontal activity and volume change in prefrontal and medial temporal lobes in aging and dementia: a review. Ageing Res Rev 2013 Mar;12(2):479-489. [doi: 10.1016/j.arr.2012.11.001] [Medline: 23183352]
40. Guarnieri RV, Alves MV, Nogueira AML, et al. Haloperidol-induced impairment on working memory capacity affecting long term memory performance: the binding hypothesis. Neuroscience. . [doi: 10.1101/2019.12.24.888073]
41. Moore T, Zirnsak M. Neural mechanisms of selective visual attention. Annu Rev Psychol 2017 Jan 3;68(1):47-72. [doi: 10.1146/annurev-psych-122414-033400] [Medline: 28051934]
42. Giuffrida V, Marc IB, Ramawat S, et al. Reward prospect affects strategic adjustments in stop signal task. Front Psychol 2023;14:1125066. [doi: 10.3389/fpsyg.2023.1125066] [Medline: 37008850]
43. Willis SL, Tennstedt SL, Marsiske M, et al. Long-term effects of cognitive training on everyday functional outcomes in older adults. JAMA 2006 Dec 20;296(23):2805-2814. [doi: 10.1001/jama.296.23.2805] [Medline: 17179457]
44. Ball K, Berch DB, Helmers KF, et al. Effects of cognitive training interventions with older adults: a randomized controlled trial. JAMA 2002 Nov 13;288(18):2271-2281. [doi: 10.1001/jama.288.18.2271] [Medline: 12425704]
45. Belchior P, Marsiske M, Sisco SM, et al. Video game training to improve selective visual attention in older adults. Comput Human Behav 2013 Jul 1;29(4):1318-1324. [doi: 10.1016/j.chb.2013.01.034] [Medline: 24003265]
46. Ferreira S, Marmeleira J, Del Pozo-Cruz J, et al. Acute effects of augmented reality exergames versus cycle ergometer on reaction time, visual attention, and verbal fluency in community older adults. Int J Environ Res Public Health 2022 Nov 8;19(22):14667. [doi: 10.3390/ijerph192214667] [Medline: 36429385]
47. Chambon C, Herrera C, Romaiguere P, Paban V, Alescio-Lautier B. Benefits of computer-based memory and attention training in healthy older adults. Psychol Aging 2014 Sep;29(3):731-743. [doi: 10.1037/a0037477] [Medline: 25244490]
48. Isbel B, Weber J, Lagopoulos J, Stefanidis K, Anderson H, Summers MJ. Neural changes in early visual processing after 6 months of mindfulness training in older adults. Sci Rep 2020 Dec 3;10(1):21163. [doi: 10.1038/s41598-020-78343-w] [Medline: 33273707]
49. Giles GM, Radomski MV, Wolf TJ. Cognition, cognitive rehabilitation, and occupational performance. Am J Occup Ther 2019 Nov 1;73(Supplement_2):7312410010p1-7312410010p25. [doi: 10.5014/ajot.2019.73S201]
50. Giles GM, Edwards DF, Morrison MT, Baum C, Wolf TJ. Screening for functional cognition in postacute care and the Improving Medicare Post-Acute Care Transformation (IMPACT) Act of 2014. Am J Occup Ther 2017;71(5):7105090010p1-7105090010p6. [doi: 10.5014/ajot.2017.715001] [Medline: 28809645]
51. Giles GM, Edwards DF, Baum C, et al. Making functional cognition a professional priority. Am J Occup Ther 2020;74(1):7401090010p1-7401090010p6. [doi: 10.5014/ajot.2020.741002] [Medline: 32078504]
52. Giles GM, Goverover Y, Foster ER, et al. Functional cognition: moving the field forward. Occup Ther J Res ;2024:15394492241296811. [doi: 10.1177/15394492241296811]
53. Wesson J, Clemson L, Brodaty H, Reppermund S. Estimating functional cognition in older adults using observational assessments of task performance in complex everyday activities: A systematic review and evaluation of measurement properties. Neuroscience & Biobehavioral Reviews 2016 Sep;68:335-360. [doi: 10.1016/j.neubiorev.2016.05.024]
54. Greenwood PM, Lambert C, Sunderland T, Parasuraman R. Effects of apolipoprotein E genotype on spatial attention, working memory, and their interaction in healthy, middle-aged adults: results From the National Institute of Mental Health's BIOCARD study. Neuropsychology 2005 Mar;19(2):199-211. [doi: 10.1037/0894-4105.19.2.199] [Medline: 15769204]
55. Kim YS, Park JH, Lee SA. Is a program to improve grocery-shopping skills clinically effective in improving executive function and instrumental activities of daily living of patients with schizophrenia? Asian J Psychiatr 2020 Feb;48:101896. [doi: 10.1016/j.ajp.2019.101896]
56. Mouri N, Ohta R, Sano C. Effects of shopping rehabilitation on older people's daily activities. Int J Environ Res Public Health 2022 Jan 5;19(1):569. [doi: 10.3390/ijerph19010569] [Medline: 35010839]
57. Okahashi S, Seki K, Nagano A, Luo Z, Kojima M, Futaki T. A virtual shopping test for realistic assessment of cognitive function. J Neuroeng Rehabil 2013 Jun 18;10(1):1-13. [doi: 10.1186/1743-0003-10-59] [Medline: 23777412]
58. Barnett MD, Hardesty DR, Griffin RA, Parsons TD. Performance on a virtual environment shopping task and adaptive functioning among older adults. J Clin Exp Neuropsychol 2023 Aug;45(5):464-472. [doi: 10.1080/13803395.2023.2249175] [Medline: 37638858]
59. Ghazal M, Yaghi M, Gad A, et al. AI-powered service robotics for independent shopping experiences by elderly and disabled people. Appl Sci (Basel) 2021;11(19):9007. [doi: 10.3390/app11199007]
60. Merritt BK, Fisher AG. Gender differences in the performance of activities of daily living. Arch Phys Med Rehabil 2003 Dec;84(12):1872-1877. [doi: 10.1016/s0003-9993(03)00483-0] [Medline: 14669197]
61. Sheehan CM, Tucker-Drob EM. Gendered expectations distort male–female differences in instrumental activities of daily living in later adulthood. J Gerontol B Psychol Sci Soc Sci 2019 Apr 12;74(4):715-723. [doi: 10.1093/geronb/gbw209] [Medline: 28158847]

Abbreviations

ASPE HHS: Office of the Assistant Secretary for Planning and Evaluation at the US Department of Health and Human Services
CS-C: correlation stability coefficient
EBIC: Extended Bayesian Information Criterion
FPL: Federal Poverty Level
IADL: instrumental activity of daily living
LASSO: least absolute shrinkage and selection operator
NHATS: National Health and Aging Trends Study

Edited by Y Jiang; submitted 16.10.24; peer-reviewed by D Fong, K Nanjo, Anonymous, Y Lee; revised version received 17.02.25; accepted 19.02.25; published 19.03.25.

Please cite as:
Li J, He R, Hsu EC, Li J
Network Analysis of Key Instrumental Activities of Daily Living and Cognitive Domains for Targeted Intervention in US Older Adults Without Dementia: Cross-Sectional Study
JMIR Aging 2025;8:e67632
URL: https://aging.jmir.org/2025/1/e67632
doi:10.2196/67632

Original Paper

Online Group–Based Dual-Task Training to Improve Cognitive Function of Community-Dwelling Older Adults: Randomized Controlled Feasibility Study

Pui Hing Chau[1], PhD; Denise Shuk Ting Cheung[1], PhD; Jojo Yan Yan Kwok[1], PhD; Wai Chi Chan[2], MBChB; Doris Sau Fung Yu[1], PhD

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

[2]Department of Psychiatry, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Pui Hing Chau, PhD
School of Nursing
The University of Hong Kong
5/F, 3 Sassoon Road
Pokfulam
Hong Kong
China (Hong Kong)
Phone: 852 39176626
Email: phchau@graduate.hku.hk

Abstract

Background: Cognitive training for older adults is crucial before cognitive impairment emerges. During periods of social distancing like the COVID-19 pandemic, cognitive stimuli are lacking. Online dual-task training is proposed as a solution to address these needs.

Objective: We aimed to explore the feasibility, acceptance, and potential effects of online group-based dual-task training as an intervention for enhancing cognitive function among community-dwelling older adults.

Methods: A randomized controlled feasibility study was conducted with 76 participants in Hong Kong, randomly assigned to the intervention and attention control groups in a ratio of 2:1 (n=50, 66% and n=26, 34%, respectively). The intervention group underwent 60-minute online dual-task training sessions twice a week for 12 weeks, incorporating cognitive components (upper limb and finger movement, arithmetic operation, and verbal fluency) and physical components (chair-based exercises) developed through a co-design approach. The attention control group received online health talks. Outcomes related to feasibility and acceptance included class attendance and self-reported satisfaction. Main outcomes related to potential effects included the Memory Inventory in Chinese and the Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) at baseline, 6 weeks (midintervention), 12 weeks (postintervention) and 18 weeks (follow-up). Descriptive statistics and linear mixed effects models were used. Effect size was described with Cohen *d*. Qualitative feedback was collected from 12 informants and analyzed by thematic analysis.

Results: About 72% (36/50) of the participants in the intervention group and 62% (16/26) in the control group attended over 75% of the classes. In total, 44 (88%) participants from the intervention group provided acceptance feedback; 82% (36/44) were satisfied and 84% (37/44) would recommend the training to others. Improvement in the Memory Inventory in Chinese score in the intervention group was observed at midintervention, postintervention, and follow-up, with a medium-to-large effect size (*d*=0.65, 0.43 and 0.85, respectively). Adjusting for baseline values, the between-group differences in the Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) score attained a small-to-medium effect size at midintervention (*d*=0.34) and postintervention (*d*=0.23). Qualitative feedback highlighted the timesaving and convenient aspects of online dual-task training, with participants finding the sessions challenging and enjoyable, and reporting benefits across cognitive, physical, and psychosocial domains. However, a preference for traditional in-person training was noted among the older adults despite the advantages of online training.

Conclusions: Online dual-task training is a feasible intervention accepted by the older adults, with potential benefits in cognitive abilities. Online training may complement in-person sessions. Further investigation in a full-scale randomized controlled trial is warranted to comprehensively explore its effects and address areas for improvement.

Trial Registration: ClinicalTrials.gov NCT05573646; https://clinicaltrials.gov/study/NCT05573646

(JMIR Aging 2025;8:e67267) doi:10.2196/67267

KEYWORDS

cognitive training; dual-task; co-design; online; older adults; Hong Kong

Introduction

Cognitive Decline

The cognitive health of older people is critical to aging in place. Even among healthy older adults, about one-third may experience cognitive decline within 18 months [1], impacting decision-making and daily activities like managing finances, medications, transportation, and meal preparation [2]. Cognitive decline is also linked to decreased physical performance and psychosocial issues, such as loneliness and depression [3]. The health care burden may thus increase in the long run [3].

Existing Dual-Task Interventions

Interventions to enhance cognitive function or to prevent decline are essential for healthy aging. When older adults have limited capacity in performing both functional tasks and cognitive tasks, they tend to prioritize the former over the latter because of the lower chance of being injured [4]. Dual-task training, which involves combining cognitive and physical training components, has been shown to improve physical and cognitive performance, including, but not limited to, gait, balance, and memory and cognitive function [5]. Compared with training with single components, dual-task training tends to have potentially greater benefits [6,7]. Dual-task training has been widely applied among populations with clinical conditions, such as stroke, Parkinson disease, and cognitive impairment. A systematic review of 13 randomized controlled trials (RCTs) on stroke survivors reported that dual-task training could improve walking and balance function [8]. A recent meta-analysis of 14 RCTs estimated a small-to-moderate effect in improving cognitive function among stroke survivors [9]. Meanwhile, a meta-analysis of 17 RCTs on people with Parkinson disease revealed positive effects in physical performance, such as gait and balance, with effect size varying from small to large [10], and a narrative review of 3 studies reported potential improvements in cognitive function among this group, although statistical significance was not always achieved [11]. A 2022 network meta-analysis concluded that dual-task training was promising for addressing the cognitive and motor symptoms of patients with Parkinson disease [12]. The benefits of dual-task training for cognitive function were commonly investigated among cognitively impaired people. A review of 18 RCTs on those with dementia or mild cognitive impairment reported improvements in overall cognitive function, attention, and functional mobility [7]. Another meta-analysis of 21 RCTs reported small-to-moderate effects in global cognitive function, memory, executive function, and attention in cognitively impaired people [13]. A recent meta-analysis of 20 RCTs also reported benefits for global cognition, executive function, and working memory with effect sizes that were moderate, small to moderate, and moderate to large, respectively [14]. There is growing literature on studies targeting cognitively healthy older people. A meta-analysis of 8 controlled trials on cognitively healthy older adults reported improvement in cognitive functions, such as global cognition, working memory, and executive function [15]. The same review also reported no statistically significant difference in efficacy between cognitively healthy older adults and those with mild cognitive impairment [15].

Service Gaps

In practice, motivating healthy older adults to engage in cognitive training can be challenging, especially when they perceive no immediate risk of decline. Presenting interventions as recreational and stimulating can enhance participation. Compared with solely cognitive training, dual-task training may address this need. Meanwhile, when there was social distancing, such as during the COVID-19 pandemic, older adults lost their opportunity to exercise and interact with others. Reductions in physical activity and social interaction not only challenged their physical and psychological health but also challenged their cognitive performance. Although group-based dual-task training could be beneficial, existing dual-task training programs have been conducted face-to-face, which limits their application during times of social distancing, prompting the exploration of online delivery methods. As the participants could join the training in their own homes, the training components had to be specially designed to simultaneously ensure safety, enjoyment, and efficacy. Hence, as an initial step, the feasibility, acceptability, and potential effects of the proposed intervention had to be investigated through a feasibility study.

Aim and Objectives

This study aimed to explore the feasibility, acceptance, and potential effects of online group-based dual-task training as an intervention for enhancing cognitive function among community-dwelling older adults. To accomplish this aim, we had three objectives: (1) to develop the intervention through a co-design approach, (2) to explore its feasibility and acceptance among older adults, and (3) to examine its potential effects in terms of the cognitive, physical, and psychosocial status of the intervention group before and after the intervention and in comparison with the control group. We hypothesized that online based dual-task training would be feasible for healthy older adults.

Methods

Study Design

This was a randomized controlled feasibility study with 2 parallel arms. The trial was registered at ClinicalTrials.gov (NCT05573646).

Ethical Considerations

Ethics approval was obtained from The University of Hong Kong/Hospital Authority Hong Kong West Cluster Institutional Review Board (UW22-038).

Setting

This study was conducted in both online and community settings across Hong Kong. Data collection was conducted at older adults community centers, an older adult educational center, and a university campus, and the intervention was delivered online.

Target Population and Sample

The target population comprised community-dwelling older adults. The inclusion criteria were (1) age ≥65 years, (2) no communication problems, and (3) ability to use an online meeting platform. The exclusion criteria were (1) contraindications to chair-based exercises and (2) engagement in any kind of cognitive training 3 months before the study or during the study period.

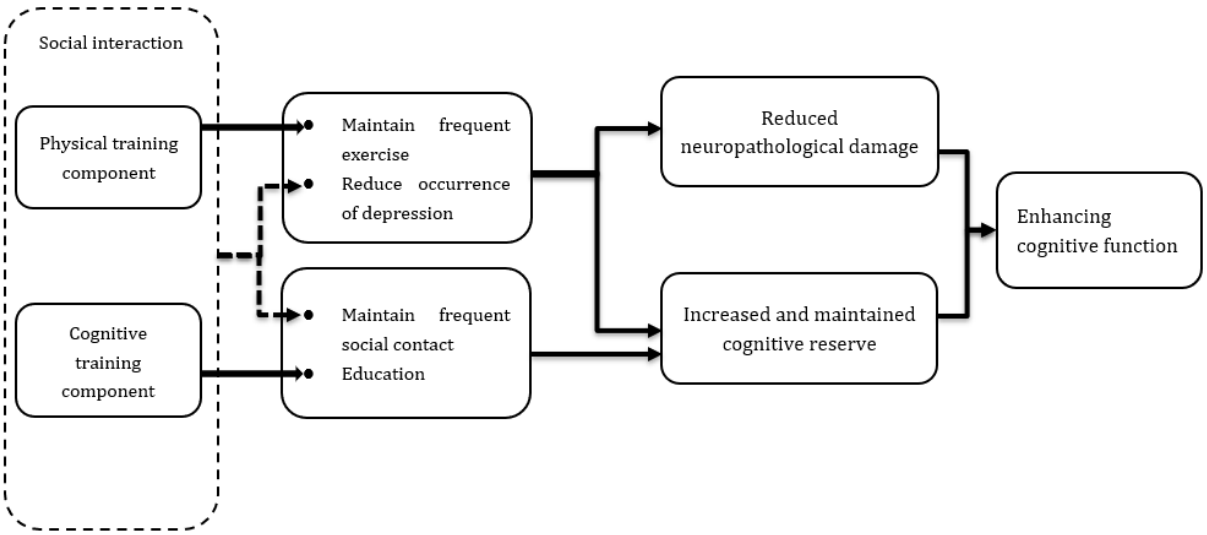
According to the literature, for a main trial designed with 80% power and 2-sided 5% significance to detect a small-to-medium effect (Cohen *d* between 0.1 and 0.3), a sample size of at least 20 per treatment arm is needed for a pilot study [16]. Meanwhile, to allow the sample to detect a medium within-group effect size in the intervention group (d=0.5) with 80% power and 2-sided 5% significance, a sample size of 34 was required for the intervention group. Taking both scenarios and 20% attrition into account, 50 participants in the intervention group and 25 participants in the control group were required.

Intervention: Online Group-Based Dual-Task Training

Participants in the intervention group received a 1-hour online group-based dual-task training session twice a week for 12 weeks, resulting in a total intervention time of 24 hours. Such frequency and duration were within the range of 4 to 25 weeks and 30 to 240 minutes per week, as used in previous dual-task training programs for healthy older adults [6].

Figure 1 shows the conceptual framework underpinning the development of the intervention based on the 2020 Report of the Lancet Commission on dementia prevention, intervention, and care [17]. The framework proposed that modifiable risk factors prevent dementia through reduction of neuropathological damage and increase and maintenance of a cognitive reserve. Among these factors, maintenance of frequent exercise and reduction of depression act on both paths, while maintenance of frequent social contact and education act on the cognitive reserve path. In our intervention, the physical training components helped to maintain frequent exercise (at least twice weekly) and the cognitive training component mimicked the education effect. The social interaction via group Zoom (Zoom Communications, Inc) classes maintained frequent social contact for at least twice a week and thus could help to relieve depression symptoms.

Figure 1. Conceptual framework.



In the development of the cognitive training component, various factors were considered, including the lower education level of the local older adults, the online delivery format of the intervention, the cultural appeal to the local older population, and the effectiveness and safety of the training. A diverse set of training components were incorporated, such as upper limb and finger movements, arithmetic operations, and verbal fluency exercises, based on their proven efficacy in previous dual-task training studies [18,19]. A co-design approach was adopted to develop the dual-task training components. In total, 9 participants from the target population were invited to a trial run to provide feedback on their interests, challenges, and suggestions for tasks they preferred. The suggested tasks were then reviewed by the project team to assess their intensity and potential impact on training outcomes. Considering what would be acceptable and effective, 18 upper limb and finger movement games, 22 arithmetic operation games, and 6 verbal fluency training components were finalized (Textbox 1). During the 1-hour training session, participants engaged in different tasks according to a predefined schedule (Table 1). Specifically, participants spent 12 minutes on upper limb and finger movement exercises, 15 minutes on arithmetic operations, and 15 minutes on verbal fluency tasks. Each game started at the easiest level, with the facilitator adjusting the cognitive demand

as participants progressed. For each game, the facilitator would start with the easiest level first. To sustain motivation and enjoyment, the difficulty level of each component increased when participants demonstrated mastery, typically indicated by 3 consecutive rounds without losing a game.

Textbox 1. Examples of cognitive training components.

Upper limb and finger movements

- Goal-directed finger movements, such as counting and reverse counting with the fingers and making motions in different directions with different hands
- Forming patterns in the air

Arithmetic operation

- Counting
- Reverse counting
- Counting or reverse counting with subtraction or addition
- Clapping at specific numbers or multiples

Verbal fluency

- "Jie Long" (a well-known Chinese word chain game)
- Naming games (particularly on Chinese cultural concepts, such as Chinese dim sum and greetings in the Chinese new year)
- Memory games (participants needed to recap the previous terms presented by others and then add their own)

Table 1. Dual-task training schedule for each 60-minute session.

Time (minutes)	Dual-task training	
	Physical components (chair-based)	Cognitive components
5	Warm-up exercise	Safety reminder or casual chit-chat
3	Toes raises	Upper limb and finger movement 1
3	Heel raises	Upper limb and finger movement 2
3	Toes raises	Upper limb and finger movement 3
3	Heel raises	Upper limb and finger movement 4
3	Break	Break
3	Stepping	Arithmetic operation 1
3	Rest	Arithmetic operation 2
3	Stepping	Arithmetic operation 3
3	Rest	Arithmetic operation 4
3	Stepping	Arithmetic operation 5
3	Break	Break
3	Stepping	Verbal fluency 1
3	Clam exercise	Verbal fluency 2
3	Stepping	Verbal fluency 3
3	Clam exercise	Verbal fluency 4
3	Stepping	Verbal fluency 5
3	Clam exercise	Verbal fluency 6
4	Cool-down exercise	Safety reminder

For the physical training component, we selected chair-based exercises, which were chosen due to their suitability for the limited space available in most local households. Chair-based exercises offer various benefits, such as improving mood and well-being, enhancing certain activities of daily living, promoting social interaction, and increasing muscle strength [20]. Alongside the cognitive training games, participants engaged in exercises like toe or heel raises on alternate sides,

stepping, and clam exercises (Table 1). The exercises were performed at a low frequency, such as 60 steps per minute. To ensure safety, participants were permitted to use their hands for balance support during stepping or clam exercises. Therefore, when upper limb and finger movement tasks were involved, only toe or heels raise were performed.

A group-based training approach was adopted as group-based cognitive training might yield greater interaction and cognitive benefits than individual-based training [21]. Group-based exercise was also reported as having the additional advantage of creating a social support environment that could enhance psychosocial well-being [22]. Moreover, group-based interventions tend to have higher adherence rates than general exercise programs due to the motivational support provided by group dynamics [23]. To encourage interaction among participants and enable close monitoring by the facilitator, a group size of less than 10 people was used.

Zoom was selected as the training platform due to its widespread use and familiarity among local older adults. Throughout the session, participants were instructed to sit on a sturdy chair with a back rest in front of a desktop computer, laptop, tablet, or mobile phone securely placed on a table. Mobile device users were provided with stands for better posture (all participants received the stand as a souvenir). To allow enough space for movement, the chair and table needed to be positioned 60 cm apart. Participants were encouraged to turn on their cameras during the session to facilitate interaction with one another. They were also asked to keep their audio on to actively participate in the cognitive tasks by verbalizing their answers.

The sessions were led by a trained facilitator. The facilitator did not need to be health professional, but they were required to undergo necessary training to ensure they had the knowledge and skills to effectively work with older adults [20]. A written manual with all the rules and regulations of the training was provided. During the initial sessions of the intervention, the project lead provided on-site supervision of the facilitator to ensure that the intervention was delivered according to the protocol. In addition, random spot checks were conducted by the project lead to ensure strict adherence to the protocol.

Attention Control: Interactive Health Talks

In line with recommendations for attention control groups [24], participants in the control group underwent 8 one-hour online group-based interactive health talks on food label knowledge and its practical applications. These sessions were led by a facilitator with nutrition training.

Outcome Measures

The primary outcomes were related to feasibility and acceptance. Feasibility was reflected by class attendance. Reasons for absence were documented when available. There was no consensus on the definition of high attendance [25]. We took 75% attendance as high attendance, as a review reported 3 of 11 studies on exercise interventions adopted this definition [25]. The proportion of participants with high attendance was calculated using the number of participants randomized to the group as denominator. A review of group exercise interventions for older adults reported that in 4 of 6 studies, 65% to 67% of participants adhered to the program [22]. Hence, our study considered over 65% of participants completing at least 75% of the classes as feasible.

Acceptance was reflected by 5-point Likert scale questions on the level of satisfaction with the intervention (“Are you satisfied with this training?”) and the level of likeliness of recommending the intervention to other people (“Would you recommend this training to friends and family”), with answers collected at the end of the intervention via instant text messaging. The proportion of respondents providing a positive response was calculated. As a conservative estimate, another proportion was calculated by assuming those who did not provide feedback as not being satisfied and not likely to recommend the intervention, with the number of participants randomized to the group as the denominator.

The secondary outcomes were related to potential effects, which included subjective memory complaints, cognitive status, working memory, executive function, physical function, instrumental activities of daily living, happiness, and social networks.

Subjective memory complaints were assessed using the Memory Inventory in Chinese (MIC) [26]. This scale consists of 27 items on memory concerns related to daily activities in the past month. Responses are rated on a scale from 0 (none) to 4 (once or more per day or continuously), with scores totaling between 0 and 108. A higher score indicates a greater frequency of memory concerns.

Cognitive status was assessed using the validated Montreal Cognitive Assessment 5 Minutes (Hong Kong Version) (HK-MoCA 5-Min), which covers 4 domains, namely, attention, executive functions or language, orientation, and memory [27]. The total score ranges from 0 to 30, with a higher score indicating better cognitive status.

Working memory was assessed by Digit Span Test [28]. Participants were required to repeat progressively longer series of digits. The test began with 2 digits and increased in length, with a maximum of 8 digits for the forward test and 7 digits for the backward test. The participant’s ability to recall the longest digit series in forward and backward order was recorded as the forward and backward scores, respectively. A higher score indicated a better working memory capacity.

Executive function was assessed by the Chinese version of the Victoria Stroop Test [29]. The test consisted of 3 parts, each presenting different stimuli: colored dots, common words unrelated to color, and color words. Participants were required to name the color in which the stimuli were printed. Inference scores were calculated based on the difference in completion time between the word or color test and the dot test, with a longer completion time indicating a poorer condition.

Lower-limb muscle strength was assessed by the 5-time chair stand task [30]. Participants were timed on how long it took them to complete the task, with a longer time indicating poorer performance. If participants were unable to finish the task within 1 minute, the test was stopped, and a time of 60 seconds was imputed for analysis purposes.

Instrumental activities of daily living were assessed by Lawton's 8 selected tasks, which include telephone use, shopping, food preparation, housekeeping, laundry, transportation, medicine use, and handling finances [31]. Participants rated their level of dependence in performing each task on a 5-point Likert scale. The score was the sum of all items, ranging from 0 to 8, with a higher score indicating greater independence.

Happiness was assessed by the Chinese version of the validated 4-item Subjective Happiness Scale [32]. Participants rated their responses on a 7-point Likert scale. The score was the average of the 4 items, ranging from 1 to 7, with a higher score indicating a higher level of happiness.

Depressive symptoms were assessed by the Patient Health Questionnaire-9 (PHQ-9) [33]. PHQ-9 was validated as a reliable tool for evaluating depression among older adults. The scale ranged from 0 to 27, with a higher score indicating a higher level of depressive symptoms.

Social networks were assessed by the Chinese version of the 6-item Lubben Social Network Scale [34]. Participants rated their responses on a 5-point Likert scale. The total score ranged from 0 to 30, with a higher score indicating a stronger social network.

All effect-related outcomes were assessed at recruitment (T0, baseline), 6 weeks after baseline (T1, midintervention assessment), 12 weeks after baseline (T2, postintervention assessment), and 18 weeks after baseline (T3, follow-up assessment). Demographic, medical and lifestyle information, such as age, gender, education level, exercise habits (measured by the International Physical Activity Questionnaire-Short Form), frailty status (measured by the FRAIL scale, which consists of 5 items, namely, fatigue, resistance, ambulation, illnesses, and loss of weight) and comorbidity (coexistence of 2 or more chronic illnesses) were collected at baseline. For participants who could not undergo face-to-face assessment for reasons such as quarantine or sickness, the 5-time chair stand task was not administered and was regarded as missing data.

Qualitative feedback was collected from a purposive sample at the end of the follow-up period through phone interviews. A semistructured interview guide was used to explore reasons for participating in the intervention, perceived difficulties encountered, logistical aspects of the intervention, and strengths and limitations perceived by participants. Prompts were provided as appropriate. Participants who exhibited varying degrees of change in the MIC score were invited to ensure diverse perspectives. Data saturation was achieved with 12 informants.

Procedure

Participants were recruited through social media and nongovernmental organizations in Hong Kong. They were randomly assigned to the intervention or control groups at a ratio of 2:1. As a strategy to enhance recruitment, a ratio of 2:1 was proposed for the number of people in the intervention and control groups, such that less participants would be disappointed with the allocation to the control group. A face-to-face orientation session was conducted to guide participants on using the Zoom platform and provide them with mobile device stands. Furthermore, for intervention group participants, demonstrations of the chair-based exercises were provided as they were unable to observe lower-limb movements on Zoom.

Randomization and Concealment

Block randomization with varying block sizes was used to generate the allocation sequence. An independent research assistant generated the sequence using an online platform and prepared sequentially numbered opaque sealed envelopes. These envelopes were opened in front of the participants immediately after obtaining informed consent.

Blinding

This was a single-blinded study as participants and interventionists could not be blinded. Trained research assistants involved in assessing outcomes related to potential effects were blinded to the allocation.

Data Analysis

The participants' characteristics were summarized using descriptive statistics. Attendance statistics, regarding the percentage of scheduled classes attended, were calculated. The percentage of participants satisfied with the intervention and those who would recommend it to others were also calculated. Adopting intent-to-treat, linear mixed effects models were used to analyze effect-related outcomes, adjusting for sociodemographic characteristics, medical history, and lifestyle patterns. A significance level of 5% was used. However, given the small sample size, focus was placed on effect size with the corresponding 95% CI. A Cohen *d* of 0.2 was considered clinically meaningful [35]. Missing outcome measures were not imputed as the linear mixed effects models could handle missing data. Statistical analyses were performed using SPSS version 28.

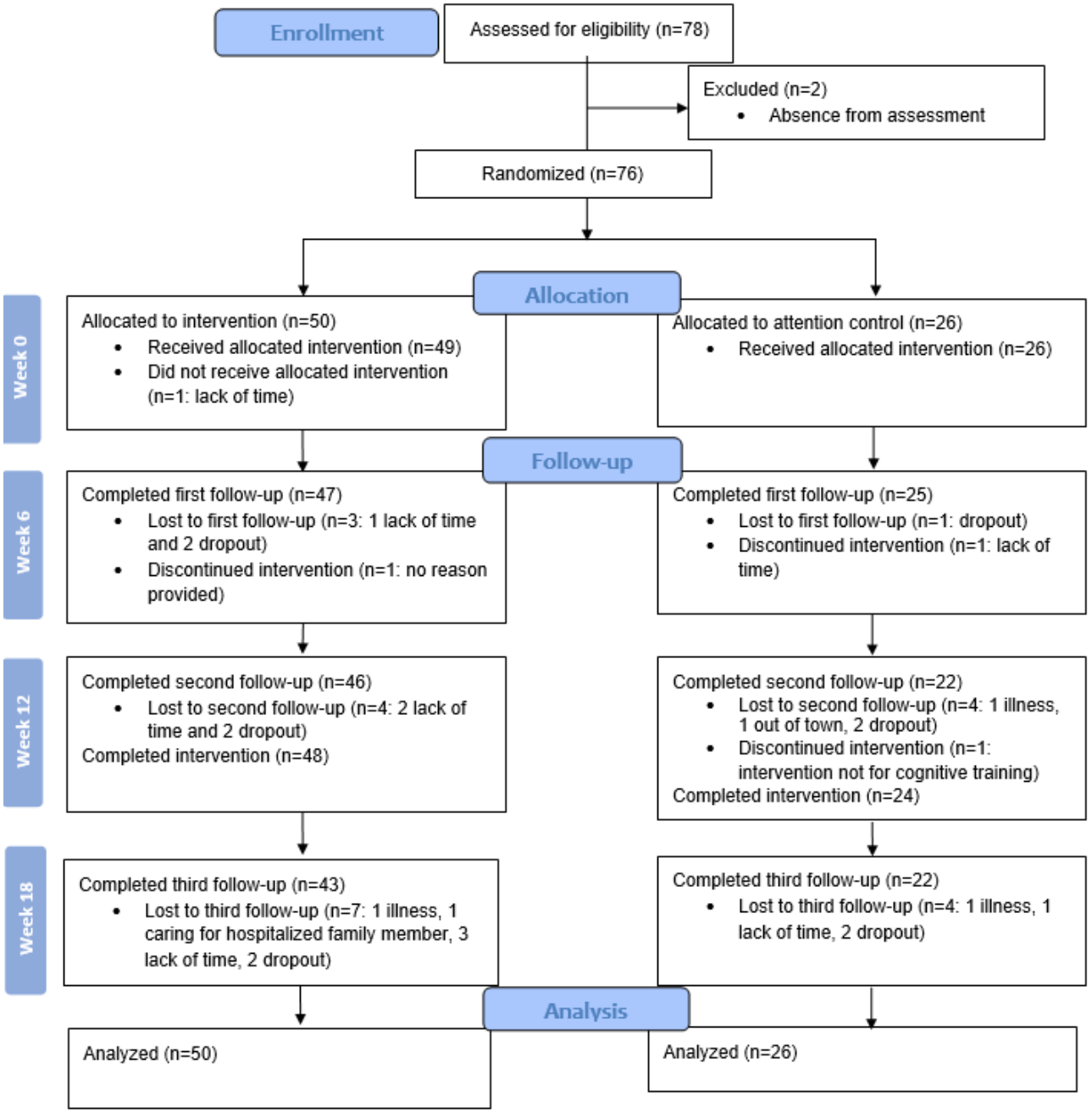
As for the qualitative interview, the audio recordings were transcribed verbatim for thematic analysis [36]. First, 2 researchers read the transcripts several times and conducted the systematic coding independently. They proceeded to generate the initial codes and the recurrent pattern within the data and identify and name the themes and subthemes according to the underlying meaning. Then, a third researcher checked for the consensus on the themes and subthemes and discussed with the team accordingly. They also worked together to identify the common threads that extend across the interviews of the participants. Analyses were performed manually without the aid of software.

Results

Overview

In total, 76 participants were recruited from October 2022 to May 2023 (Figure 2). To manage class sizes, the 50 participants in the intervention group were divided into 8 subgroups, each with 24 classes, and the 26 participants in the attention control group were divided into 2 subgroups, each with 8 classes. Overall, 48 (96%) participants from the intervention group and 24 (92%) from the control group completed the assigned intervention.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram.



The mean ages of the participants in the intervention and the control groups were 71.5 (SD 5.2) and 71.1 (SD 6.0) years, respectively. There were 64% (32/50) female participants in the intervention group and 88% (23/26) in the control group. One-fifth (15/76, 20%) of the participants had a primary level of education or below. About 44% (22/50) of the participants in the intervention group and 54% (14/26) in the control group had multimorbidity. Most participants in both groups (30/50, 60% and 16/26, 61%) had low levels of physical activity. Prefrailty was observed in about 38% (19/50) of intervention-group participants and 50% (13/26) of control-group participants. Table 2 shows the characteristics of the participants.

Table 2. Baseline characteristics of participants (N=76).

Characteristics	Intervention (n=50)	Control (n=26)
Age (y), mean (SD)	71.5 (5.2)	71.1 (6.0)
Sex, n (%)		
Male	18 (36)	3 (11)
Female	32 (64)	23 (89)
Education level, n (%)		
Primary or below	10 (20)	5 (19)
Secondary	27 (54)	14 (54)
Tertiary or above	13 (26)	7 (27)
Multimorbidity[a], n (%)		
No	28 (56)	12 (46)
Yes	22 (44)	14 (54)
Physical activity level (measured by IPAQ[b]), n (%)		
Low	30 (60)	16 (61)
Moderate	15 (30)	9 (35)
High	5 (10)	1 (4)
Frailty status (measured by FRAIL[c] scale), n (%)		
Robust	31 (62)	13 (50)
Prefrail	19 (38)	13 (50)

[a]Multimorbidity is defined as the co-existence of at least 2 of the following long-term health conditions: hypertension, heart disease, high cholesterol, cancer, osteoporosis, and other chronic illnesses, as reported by the participants.

[b]IPAQ: International Physical Activity Questionnaire.

[c]FRAIL: Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight.

Feasibility

In the intervention group, 36 (72%) participants attended over 75% of the 24 classes. For the attention control group, 16 participants (61%) attended over 75% of the 8 classes. Common reasons for absence in both groups included time clashes with medical appointments, illness, or other prior commitments.

Acceptability

Among 44 participants who provided acceptance feedback on the intervention, 82% (36/44) expressed being very satisfied or satisfied with the intervention, and 84% (37/44) indicated they would be very likely or likely to recommend the intervention to others. Assuming that those who did not provide feedback were not satisfied and were not likely to recommend the intervention, the satisfaction rate was 72% (36/50) and the recommendation rate was 74% (37/50).

Potential Effects

Regarding the effect-related outcomes, owing to the pilot nature of the trial, statistical significance was not the primary focus, instead, the effect size was noted. Within-group changes were explored and are presented in Multimedia Appendix 1. For the intervention group, the effect size of the MIC improvements at midintervention, postintervention, and follow-up compared to baseline were 0.65, 0.43 and 0.85, whereas the estimates for the control group were 0.51, 0.37 and 0.78, respectively. Similarly, both groups experienced an increase in HK-MoCA 5-Min score with moderate to large effect sizes. However, changes in working memory, as assessed by the Digit Span Test, did not consistently show the same direction in both groups. The intervention group showed a moderate improvement in executive function, as shown by the inference score for color from the Victoria Stroop Test, while the control group showed a small-to-large improvement in the inference score for words. Improvement was not observed for other outcomes in both groups.

Between-group differences, adjusting for the baseline differences, were explored and are presented in Multimedia Appendix 1. The intervention group did not outperform the control group in the MIC, but it showed better performance in the HK-MoCA 5-Min at the midintervention and the postintervention assessments (effect size 0.34 and 0.23, respectively). Meanwhile, a medium effect size was observed for between-group difference in the Digit Span Test at the follow-up. However, the findings for the Victoria Stroop Test were inconsistent. Details on other outcomes are presented in Multimedia Appendix 1.

Comparing the effect-related outcomes between participants with high attendance (≥75% of class) and low attendance in the intervention group (Multimedia Appendix 1), those with high attendance had fewer subjective memory complaints than those

with low attendance at midintervention and postintervention, with effect sizes of 1.18 and 1.53, respectively. At the follow-up assessment, the between-group difference was observed to be in favor of the high-attendance group in the Digit Span Test, inference score for color, and PHQ-9 score.

Qualitative Feedback

Table 3 shows the demographic characteristics of the 12 informants. Table 4 shows the 5 identified themes and the corresponding subthemes derived from the qualitative feedback, along with the significant quotes. The themes are summarized as presented in Textbox 2.

Table 3. Characteristics of the 12 informants providing qualitative feedback.

Informant	Sex	Age group (y)	Education level
A	Female	65-74	Primary
B	Female	65-74	Secondary
C	Male	≥75	Primary
D	Female	65-74	Secondary
E	Female	65-74	Secondary
F	Female	65-74	Tertiary
G	Female	≥75	Secondary
H	Female	65-74	Secondary
I	Female	65-74	Secondary
J	Male	65-74	Tertiary
K	Male	65-74	Primary
L	Male	65-74	Tertiary

Table 4. Five identified themes and the corresponding subthemes from the qualitative feedback and the significant quotes.

Themes and subthemes	Significant quotes by informants
Perceived strength of the training (before participation)	
Arousing curiosity and motivation to join	• "I heard an introduction at the elderly community center, and I want to learn more." [Informant A] • "Curious about online dual task training." [Informant B]
Giving them hope to slow down cognitive decline	• "I am taking precautions because I don't know if I will have cognitive impairment in the future. I just want to prevent it. Early prevention is better." [Informant D] • "I want to maintain, not improve, cognitive ability." [Informant L]
Offering opportunity for social interaction	• "I want to make new friends." [Informant C]
Save travel time	• "No need to spend time going back and forth; you can stay at home. Nowadays, we mostly go to centers to learn things. In-person is much better, but it takes more time to get there and back." [Informant E] • "But overall, I think online is good for participants, as it saves travel time." [Informant J]
Positive experience from the training	
Comprehensive training	• "I think the course is quite good and comprehensive; it does not just focus on those with cognitive impairment." [Informant I]
Challenging training	• "There is some difficulty; sometimes I forget, like moving a hand or getting distracted and forgetting to do the stepping. You need to stay focused to complete it." [Informant B] • "At the beginning, I feel a bit frustrated when playing. However, it has its benefits. If everything were easy, there would be no need for training." [Informant G] • "If you have to move both your foot and hand, you often end up neglecting one. Sometimes, you also have to keep track of numbers, so you need to pay attention to what the other participants are saying." [Informant K]
Engaging facilitator	• "The host guided us with patience." [Informant A] • "Overall, I think the team is very dedicated. Sometimes when we can't hear, she [the facilitator] reminds us. Or when we can't think of an answer, she gives us hints at the right time. I think the whole process is very smooth." [Informant I]
Peer support and encouragement during the training	• "We encourage each other during the training." [Informant L]
Enjoy the convenience at home environment	• "I can start the training right after breakfast. When we have breaks, we can go to the washroom and drink water at home. After the session ends, we can prepare lunch ourselves." [Informant A]
Challenges encountered with online training	
Constrained supervision and intensity of training	• "At least with Zoom, the camera cannot capture my feet, so you cannot tell if participants are doing stepping. I think in-person is ultimately more useful because you can see everyone's movements." [Informant B] • "Because in-person, you can immediately see everyone's performance." [Informant D]
Harder to follow the sequence if the training involved taking turns	• "Because it is online and most participants are women, it is hard for me to know that it is my turn unless there is a man preceding my turn. But after a few sessions, everyone starts to get familiar with each other, and we build up rapport about the sequence." [Informant J]
Less concentrated and less cordial	• "During training, some might not be able to concentrate, for example, if there are things happening at home." [Informant A] • "It's not as easy to get distracted if it is in-person." [Informant D]
Class atmosphere not comparable to in-person class	• "But in-person is more enjoyable, with more group interactions, which makes it more fun." [Informant G]
Technical issues encountered	• "The display is limited in size, and there might be some lag time." [Informant J] • "The problem is that some resource-limited families don't have Wi-Fi at home, so they have to go to the center to access it, which is less convenient." [Informant K]

Themes and subthemes	Significant quotes by informants
Perceived benefits of the training	
Improved cognitive ability (including attention and concentration)	• "Initially, I felt very unfamiliar with it, but after a few more sessions, I absorbed new knowledge and it made my mind more agile." [Informant G] • "I think I can use my brain more, as I usually don't use it much." [Informant H] • "My memory has been maintained without much decline." [Informant L]
Improved coordination and physical flexibility	• "After participating, my knowledge has broadened, and both my hands and mind have become more agile." [Informant C] • "I find that after training, the coordination of my hands and feet has become a bit more agile compared to before." [Informant L]
Improved communication skills	• "It can train an older person to handle different communication demands." [Informant J]
Promotion of positive mood and social interactions	• "Someone accompanies you. I also find it fun and enjoyable." [Informant F]
Learn new skills for self-practice in future	• "I have not experienced this kind of hand-eye coordination, but now I do. Sometimes, when I'm alone, I practise the movements." [Informant E] • "I'll go back and practise it myself." [Informant K]
Increased self-efficacy and self-confidence	• "My brain is now willing to exercise and let me do calculations." [Informant F] • "I actually feel quite happy and it gives a sense of accomplishment because you can achieve your goals and meet the training requirements, so there is a sense of success." [Informant J]
Self-explanation of less benefit from "cheating"	• "I do not feel like I have gained much. Maybe I prepare the answers in advance. For example, knowing that the next lesson will involve 'clapping on 7,' I will prepare the answers beforehand. But if it is in-person, it is harder to cheat." [Informant I]
Suggestions for improvements	
Continue the training (beyond 12 weeks)	• "I think I need to keep doing it consistently to see improvement. But I understand it is difficult if resources are lacking. Once the training stops, like now, I have forgotten a lot." [Informant K]
Booster training	• "Suggest to add online virtual courses. Each session would be half an hour, allowing students to choose their own study times." [Informant K] • "If possible, it would be best to have a booster session once a month to refresh what we have learned." [Informant L]
Consider in-person mode (but reduced frequency to accommodate time and venue constraints)	• "I think in-person sessions are ultimately more useful because you can see everyone's movements. In-person sessions could be held once a week, not too frequently, since not everyone can attend twice a week, and the center might not be able to provide the space for two days of training." [Informant B]

Textbox 2. Summary of the themes and subthemes derived from qualitative feedback.

- Perceived strength of the training (before participation): informants found dual-task training to be a novel approach, which interested them and motivated them to participate. They expected the training would help to slow down cognitive decline. The online format of the training saved them travel time but could still offer social interaction.
- Positive experience from the training: informants described the training comprehensive and challenging. The engaging facilitator increased their adherence. The group training enabled peer support, enhancing their training experience. Informants also found having the training at home very convenient.
- Challenges encountered with online training: supervision and the intensity of training was limited by the online format. Some informants found it harder to follow the sequence if the training involved taking turns. Some found the home environment made them less concentrated and the class atmosphere was not comparable to in-person class. The quality of the experience was also influenced by the display size of their device, Wi-Fi speed, and more importantly accessibility to a stable internet connection.
- Perceived benefits after the training: informants perceived benefits in cognitive ability, coordination, physical flexibility, communication skills, positive moods, and social interactions. Some valued acquiring new skills that could enable future self-practice. Some mentioned increased self-efficacy and self-confidence. Meanwhile, one mentioned the use of a cheat sheet as related to the negligible improvement.
- Suggestions for improvements: informants expressed a desire to continue the training by either extending the training period or offering regular booster sessions. However, informants generally preferred in-person training, although they acknowledged limitations regarding time and venue.

Discussion

Principal Findings

This study demonstrated feasibility and explored potential effects of an online dual-task training for community-dwelling older adults. The innovative aspects of the online dual-task training included (1) the delivery mode, allowing for uninterrupted training even during a pandemic, (2) chair-based exercises suitable for home-based practice, (3) cognitive training tailored to the needs of Chinese older adults with lower education levels, and (4) a co-design approach that considered the preferences and interests of the participants when designing the training components.

This study was launched at the end of 2022 when the COVID-19 pandemic was still ongoing. While there was no citywide quarantine policy in place, there was still a home quarantine policy for those who were infected and their close contacts. Older adults tended to limit their outings to reduce the risk of infection, as well as chance of becoming close contacts. The older population had experienced a lack of activities for a prolonged period, potentially leading to a faster decline in cognitive performance. The group-based online dual-task training addressed this gap by offering training that the participants could participate in at home.

The findings of our pilot study showed the intervention was highly feasible, as 72% (36/50) of the participants in the intervention group attended over three-quarters of the classes, which was above the target of 65%. The intervention was also well received by the participants; 82% (36/44) of the participants were satisfied with the intervention and 84% (37/44) would recommend the intervention to others. Even adopting the most conservative approach, such figures remained high at 72% (36/50) and 74 (37/50), respectively. The dual-task training successfully attracted some relatively cognitively healthy participants to join and could be continued irrespective of the pandemic situation. This initiative addressed the existing gap in services by providing online cognitive training for community-dwelling older adults.

Within-group improvement was observed in subjective memory complaints and cognitive function in the intervention group. Such improvements became evident at midintervention and persisted until the final follow-up, showing a medium-to-large effect size. Despite the qualitative feedback indicating self-perceived benefits, other outcomes did not consistently favor the intervention. Controlled for the baseline outcomes, small-to-medium effects ranging from 0.23 to 0.34 were observed in cognitive function for the between-group difference in midintervention and postintervention assessments. These findings are consistent with previous literature reporting small effects of 0.22 to 0.29 for dual-task training over control groups [37,38]. Nevertheless, inconsistent results were reported for other outcomes. Moreover, the fluctuation in the magnitude of the improvement was unexpected. Although possible self-practicing of the dual-task training after the intervention period might explain the increase in within-group improvement at the follow-up, the drop at the postintervention assessment was unexpected. There were factors that might have affected the evaluation. For example, engagement in the environment might influence achievement in relation to the effects of cognitive training [39], but the social context of the participants was not assessed. Another possibility might be COVID-19 infection, which might influence the performance of the participants not only during the infection period but afterward [40]. Moreover, some assessments could not be performed face-to-face but were conducted via Zoom; the different modes of data collection might have also influenced their performance. After the pandemic, a proper RCT should be conducted.

Meanwhile, within-group improvement was observed for the attention control group. The attention control condition consisted of a series of interactive health talks focusing on nutrition and food labeling. Participants memorized the contents delivered in each talk so they could answer the instructors' questions. Moreover, some participants had to learn to operate the Zoom platform as it was new to them. The learning process might also be beneficial to the participants, as some calculation was involved in comparing food labels using different units. In future trials, usual care control may be adopted.

Qualitative feedback highlighted the strength of the training regarding convenience, innovation, and comprehensiveness. In contrast, participants expressed a preference for in-person interactions. Hence, online training could be considered as a potential alternative or complement to traditional in-person training. Participants engaged in the training before receiving a cognitive impairment diagnosis, indicating that dual-task training could motivate early engagement in cognitive training. The facilitator leading the training does not necessarily be a health professional [20], and it is the engaging nature of the training, as revealed from the feedback, that the participants highly valued. This makes implementation of the intervention more feasible. Apart from the cognitive training, the participants also treasured social interactions and peer support during the class, and these further support the need to organize group-based training.

Implications

Dual-task training offers sufficient challenges to community-dwelling older adults; they found it interesting and reported that it boosted their motivation to participate in cognitive training. The contents of the training were comprehensive, covering motion, arithmetic operations, and verbal fluency. The training was conducted in a stimulating and recreational atmosphere, which was well received by the participants.

Online training offers the advantage of saving time and being convenient. However, it is not intended to replace traditional in-person training, as older adults generally prefer face-to-face interactions. Nonetheless, online training can be viewed as an alternative or complement to in-person sessions. Participants highlighted that organizing training twice a week in-person would be challenging, indicating a potential solution of combining one day of in-person training with one day of online training to maintain a twice-weekly training frequency. As older adults become more accustomed to online training, it can serve as a viable option during situations like a pandemic, enabling

training to continue even when older adult community centers are closed or during citywide quarantine measures.

When considering the implementation of in-person or hybrid mode training, it is essential to conduct research studies to assess their effectiveness. If in-person dual-task training is to be provided, incorporating more physical components could be beneficial. Similarly, any modified training approaches should undergo evaluation before being integrated into routine services. Future research studies could involve including a usual care control group or a waitlist control group to accurately determine the true impact of the training, building upon the improvement areas identified in this study. For instance, potential strategies for future interventions may involve implementing a 6-week intervention period, as some participants in the outcome evaluations exhibited positive effects as early as 6 weeks after the intervention; combining both in-person and online training modalities; and incorporating booster sessions. The exclusion criterion regarding engagement in any form of cognitive training 3 months before or during the study period may pose challenges once activities return to normal after the pandemic. Therefore, future trials should take a pragmatic approach, ensuring that any benefits derived from the intervention are in addition to participants' existing daily lifestyle routines.

Strengths and Limitations

The strength of this study included the integration of both quantitative and qualitative outcomes, which would reflect the feasibility and perceived benefits more holistically. The multiple outcome domains also helped to examine the potential effects more comprehensively. The multiple assessment points (4 time points) would facilitate the exploration of potential effects relative to the dosage of the intervention. The innovation components of the interventions were not restricted to the local context but could largely be applied to other regions.

Nevertheless, there were some limitations. The use of an attention control group might have limited the examination of potential effects from the intervention. Moreover, COVID-19 infection among the participants or their family members and the overall social environment during the study period were not documented, limiting the ability to control for the influences from these factors on the study outcomes. The cognitive domains assessed were limited to attention, verbal fluency, orientation, working memory, and executive function. Other cognitive domains, such as perceptual-motor control and social cognition, that might benefit from the intervention were not evaluated. User satisfaction was based on a self-developed question, and a validated scale could be considered in future. Owing to the nature of the feasibility study, the sample size was small, and a proper trial is needed in the future involving larger sample size.

Conclusions

Online dual-task training was shown to be a feasible intervention and was well received by older adults during the pandemic. Within-group improvement was observed in subjective memory complaints and cognitive function, with a medium-to-large effect size in the intervention group. Further investigation in a full-scale RCT is required to fully explore the intervention's potential effects.

Acknowledgments

This project was supported by the Health and Medical Research Fund (19200991) under the Food and Health Bureau of the government of the Hong Kong Special Administrative Region. The authors would like to thank the New Territories West Elder Academies Cluster and Hong Kong Sheng Kung Hui Welfare Council for their support and the older adults for their participation.

Authors' Contributions

All authors were responsible for conceptualization, funding acquisition, methodology, and writing−review and editing. PHC was responsible for formal analysis, project administration, and writing−original draft.

Conflicts of Interest

PHC is an Associate Editor for JMIR Aging.

Multimedia Appendix 1
Within-group and between-group differences in effect-related outcomes based on mixed effects model.
[PDF File (Adobe PDF File), 140 KB - aging_v8i1e67267_app1.pdf]

Multimedia Appendix 2
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 18050 KB - aging_v8i1e67267_app2.pdf]

References

1. Woodard JL, Seidenberg M, Nielson KA, Smith JC, Antuono P, Durgerian S, et al. Prediction of cognitive decline in healthy older adults using fMRI. J Alzheimers Dis 2010;21(3):871-885 [FREE Full text] [doi: 10.3233/JAD-2010-091693] [Medline: 20634590]

2. Salthouse T. Consequences of age-related cognitive declines. Annu Rev Psychol 2012;63:201-226 [FREE Full text] [doi: 10.1146/annurev-psych-120710-100328] [Medline: 21740223]
3. Comijs HC, Dik MG, Aartsen MJ, Deeg DJ, Jonker C. The impact of change in cognitive functioning and cognitive decline on disability, well-being, and the use of healthcare services in older persons. Results of Longitudinal Aging Study Amsterdam. Dement Geriatr Cogn Disord 2005;19(5-6):316-323. [doi: 10.1159/000084557] [Medline: 15785032]
4. Li KZ, Lindenberger U, Freund AM, Baltes PB. Walking while memorizing: age-related differences in compensatory behavior. Psychol Sci 2001 May;12(3):230-237. [doi: 10.1111/1467-9280.00341] [Medline: 11437306]
5. Pereira Oliva HN, Mansur Machado FS, Rodrigues VD, Leão LL, Monteiro-Júnior RS. The effect of dual-task training on cognition of people with different clinical conditions: an overview of systematic reviews. IBRO Rep 2020 Dec;9:24-31 [FREE Full text] [doi: 10.1016/j.ibror.2020.06.005] [Medline: 33336101]
6. Lauenroth A, Ioannidis AE, Teichmann B. Influence of combined physical and cognitive training on cognition: a systematic review. BMC Geriatr 2016 Jul 18;16:141 [FREE Full text] [doi: 10.1186/s12877-016-0315-1] [Medline: 27431673]
7. Yu D, Li X, He S, Zhu H, Lam FM, Pang MY. The effect of dual-task training on cognitive ability, physical function, and dual-task performance in people with dementia or mild cognitive impairment: a systematic review and meta-analysis. Clin Rehabil 2024 Apr;38(4):443-456. [doi: 10.1177/02692155231216621] [Medline: 38009067]
8. He Y, Yang L, Zhou J, Yao L, Pang MY. Dual-task training effects on motor and cognitive functional abilities in individuals with stroke: a systematic review. Clin Rehabil 2018 Jul 23;32(7):865-877. [doi: 10.1177/0269215518758482] [Medline: 29473480]
9. Feng QM, Chen JW. Effects of exercise combined with cognitive dual-task training on cognitive function and sustained effects in stroke patients: a systematic review and meta-analysis. Disabil Rehabil 2024 Dec 06:1-10. [doi: 10.1080/09638288.2024.2435523] [Medline: 39641346]
10. García-López H, de Los Ángeles Castillo-Pintor M, Castro-Sánchez AM, Lara-Palomo IC, Obrero-Gaitán E, Cortés-Pérez I. Efficacy of dual-task training in patients with Parkinson's disease: a systematic review with meta-analysis. Mov Disord Clin Pract 2023 Sep;10(9):1268-1284 [FREE Full text] [doi: 10.1002/mdc3.13823] [Medline: 37772294]
11. Xiao Y, Yang T, Shang H. The impact of motor-cognitive dual-task training on physical and cognitive functions in Parkinson's disease. Brain Sci 2023 Mar 03;13(3):437 [FREE Full text] [doi: 10.3390/brainsci13030437] [Medline: 36979247]
12. Kwok JY, Smith R, Chan LM, Lam LC, Fong DY, Choi EP, et al. Managing freezing of gait in Parkinson's disease: a systematic review and network meta-analysis. J Neurol 2022 Jun;269(6):3310-3324. [doi: 10.1007/s00415-022-11031-z] [Medline: 35244766]
13. Ali N, Tian H, Thabane L, Ma J, Wu H, Zhong Q, et al. The effects of dual-task training on cognitive and physical functions in older adults with cognitive impairment; a systematic review and meta-analysis. J Prev Alzheimers Dis 2022;9(2):359-370. [doi: 10.14283/jpad.2022.16] [Medline: 35543010]
14. Ye JY, Chen R, Chu H, Lin HC, Liu D, Jen HJ, et al. Dual-task training in older adults with cognitive impairment: a meta-analysis and trial sequential analysis of randomized controlled trials. Int J Nurs Stud 2024 Jul;155:104776. [doi: 10.1016/j.ijnurstu.2024.104776] [Medline: 38703695]
15. Han K, Tang Z, Bai Z, Su W, Zhang H. Effects of combined cognitive and physical intervention on enhancing cognition in older adults with and without mild cognitive impairment: a systematic review and meta-analysis. Front Aging Neurosci 2022;14:878025 [FREE Full text] [doi: 10.3389/fnagi.2022.878025] [Medline: 35928994]
16. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
17. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 08;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
18. Yamada M, Aoyama T, Tanaka B, Nagai K, Ichihashi N. Seated stepping exercise in a dual-task condition improves ambulatory function with a secondary task: a randomized controlled trial. Aging Clin Exp Res 2011;23(5-6):386-392. [doi: 10.1007/BF03337763] [Medline: 21042047]
19. Morita E, Yokoyama H, Imai D, Takeda R, Ota A, Kawai E, et al. Effects of 2-year cognitive motor dual-task training on cognitive function and motor ability in healthy elderly people: a pilot study. Brain Sci 2018 May 11;8(5):86 [FREE Full text] [doi: 10.3390/brainsci8050086] [Medline: 29751661]
20. Robinson KR, Leighton P, Logan P, Gordon AL, Anthony K, Harwood RH, et al. Developing the principles of chair based exercise for older people: a modified Delphi study. BMC Geriatr 2014 May 19;14:65 [FREE Full text] [doi: 10.1186/1471-2318-14-65] [Medline: 24884392]
21. Kelly ME, Loughrey D, Lawlor BA, Robertson IH, Walsh C, Brennan S. The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: a systematic review and meta-analysis. Ageing Res Rev 2014 May;15:28-43. [doi: 10.1016/j.arr.2014.02.004] [Medline: 24607830]
22. Farrance C, Tsofliou F, Clark C. Adherence to community based group exercise interventions for older people: a mixed-methods systematic review. Prev Med 2016 Jun;87:155-166 [FREE Full text] [doi: 10.1016/j.ypmed.2016.02.037] [Medline: 26921655]

23. Ehlers DK, Daugherty AM, Burzynska AZ, Fanning J, Awick EA, Chaddock-Heyman L, et al. Regional brain volumes moderate, but do not mediate, the effects of group-based exercise training on reductions in loneliness in older adults. Front Aging Neurosci 2017;9:110 [FREE Full text] [doi: 10.3389/fnagi.2017.00110] [Medline: 28487648]
24. Aycock DM, Hayat MJ, Helvig A, Dunbar SB, Clark PC. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 2018 Jun;41(3):320-328. [doi: 10.1002/nur.21870] [Medline: 29906317]
25. Hawley-Hague H, Horne M, Skelton DA, Todd C. Review of how we should define (and measure) adherence in studies examining older adults' participation in exercise classes. BMJ Open 2016 Jun 23;6(6):e011560 [FREE Full text] [doi: 10.1136/bmjopen-2016-011560] [Medline: 27338884]
26. Lui VW, Lam LC, Chiu HF. Validation of a memory inventory for the assessment of awareness of memory deficits in Alzheimer's disease in Chinese elderly. Int J Geriatr Psychiatry 2006 Oct;21(10):917-923. [doi: 10.1002/gps.1580] [Medline: 16927409]
27. Yeung PY, Wong LL, Chan CC, Leung JL, Yung CY. A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. Hong Kong Med J 2014 Dec;20(6):504-510 [FREE Full text] [doi: 10.12809/hkmj144219] [Medline: 25125421]
28. Leung JL, Lee GT, Lam YH, Chan RC, Wu JY. The use of the Digit Span Test in screening for cognitive impairment in acute medical inpatients. Int Psychogeriatr 2011 Dec;23(10):1569-1574 [FREE Full text] [doi: 10.1017/S1041610211000792] [Medline: 21729426]
29. Lee TM, Chan CC. Stroop interference in Chinese and English. J Clin Exp Neuropsychol 2000 Aug;22(4):465-471. [doi: 10.1076/1380-3395(200008)22:4;1-0;FT465] [Medline: 10923056]
30. Takai Y, Ohta M, Akagi R, Kanehisa H, Kawakami Y, Fukunaga T. Sit-to-stand test to evaluate knee extensor muscle size and strength in the elderly: a novel approach. J Physiol Anthropol 2009;28(3):123-128 [FREE Full text] [doi: 10.2114/jpa2.28.123] [Medline: 19483373]
31. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
32. Nan H, Ni MY, Lee PH, Tam WW, Lam TH, Leung GM, et al. Psychometric evaluation of the Chinese version of the Subjective Happiness Scale: evidence from the Hong Kong FAMILY Cohort. Int J Behav Med 2014 Aug;21(4):646-652 [FREE Full text] [doi: 10.1007/s12529-014-9389-3] [Medline: 24515396]
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
34. Chang Q, Sha F, Chan CH, Yip PS. Validation of an abbreviated version of the Lubben Social Network Scale ("LSNS-6") and its associations with suicidality among older adults in China. PLoS One 2018;13(8):e0201612 [FREE Full text] [doi: 10.1371/journal.pone.0201612] [Medline: 30071067]
35. Sedaghat AR. Understanding the minimal clinically important difference (MCID) of patient-reported outcome measures. Otolaryngol Head Neck Surg 2019 Oct;161(4):551-560. [doi: 10.1177/0194599819852604] [Medline: 31159641]
36. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
37. Zhu X, Yin S, Lang M, He R, Li J. The more the better? A meta-analysis on effects of combined cognitive and physical intervention on cognition in healthy older adults. Ageing Res Rev 2016 Nov;31:67-79. [doi: 10.1016/j.arr.2016.07.003] [Medline: 27423932]
38. Gavelin HM, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
39. Park DC, Bischof GN. The aging mind: neuroplasticity in response to cognitive training. Dialogues Clin Neurosci 2013 Mar;15(1):109-119 [FREE Full text] [doi: 10.31887/DCNS.2013.15.1/dpark] [Medline: 23576894]
40. Hampshire A, Azor A, Atchison C, Trender W, Hellyer PJ, Giunchiglia V, et al. Cognition and memory after Covid-19 in a large community sample. N Engl J Med 2024 Feb 29;390(9):806-818 [FREE Full text] [doi: 10.1056/NEJMoa2311330] [Medline: 38416429]

Abbreviations

MIC: Memory Inventory in Chinese
HK-MoCA 5-Min: Montreal Cognitive Assessment 5 Minutes (Hong Kong Version)
PHQ-9: Patient Health Questionnaire-9
RCT: randomized controlled trial

Edited by Y Luo; submitted 10.10.24; peer-reviewed by A Ogunniyi, A Iizuka; comments to author 07.01.25; revised version received 02.03.25; accepted 31.03.25; published 16.05.25.

Please cite as:
Chau PH, Cheung DST, Kwok JYY, Chan WC, Yu DSF
Online Group–Based Dual-Task Training to Improve Cognitive Function of Community-Dwelling Older Adults: Randomized Controlled Feasibility Study
JMIR Aging 2025;8:e67267
URL: https://aging.jmir.org/2025/1/e67267
doi:10.2196/67267
PMID:

Comparison of 3 Aging Metrics in Dual Declines to Capture All-Cause Dementia and Mortality Risk: Cohort Study

Anying Bai[1], PhD; Shan He[2], MD; Yu Jiang[3,4], PhD; Weihao Xu[5], MD; Zhanyi Lin[5], MD

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[2]Haikou Cadre's Sanitarium of Hainan Military Region, Haikou, China

[3]School of Health Policy and Management, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[4]School of Population Medicine and Public Health, Peking Union Medical College, Beijing, China

[5]Department of Geriatrics, Guangdong Provincial Geriatrics Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, No. 106, Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

Corresponding Author:
Weihao Xu, MD
Department of Geriatrics, Guangdong Provincial Geriatrics Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, No. 106, Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

Abstract

Background: The utility of aging metrics that incorporate cognitive and physical function is not fully understood.

Objective: We aim to compare the predictive capacities of 3 distinct aging metrics—motoric cognitive risk syndrome (MCR), physio-cognitive decline syndrome (PCDS), and cognitive frailty (CF)—for incident dementia and all-cause mortality among community-dwelling older adults.

Methods: We used longitudinal data from waves 10-15 of the Health and Retirement Study. Cox proportional hazards regression analysis was employed to evaluate the effects of MCR, PCDS, and CF on incident all-cause dementia and mortality, controlling for socioeconomic and lifestyle factors, as well as medical comorbidities. Discrimination analysis was conducted to assess and compare the predictive accuracy of the 3 aging metrics.

Results: A total of 2367 older individuals aged 65 years and older, with no baseline prevalence of dementia or disability, were ultimately included. The prevalence rates of MCR, PCDS, and CF were 5.4%, 6.3%, and 1.3%, respectively. Over a decade-long follow-up period, 341 cases of dementia and 573 deaths were recorded. All 3 metrics were predictive of incident all-cause dementia and mortality when adjusting for multiple confounders, with variations in the strength of their associations (incident dementia: MCR odds ratio [OR] 1.90, 95% CI 1.30 - 2.78; CF 5.06, 95% CI 2.87 - 8.92; PCDS 3.35, 95% CI 2.44 - 4.58; mortality: MCR 1.60, 95% CI 1.17 - 2.19; CF 3.26, 95% CI 1.99 - 5.33; and PCDS 1.58, 95% CI 1.17 - 2.13). The C-index indicated that PCDS and MCR had the highest discriminatory accuracy for all-cause dementia and mortality, respectively.

Conclusions: Despite the inherent differences among the aging metrics that integrate cognitive and physical functions, they consistently identified risks of dementia and mortality. This underscores the importance of implementing targeted preventive strategies and intervention programs based on these metrics to enhance the overall quality of life and reduce premature deaths in aging populations.

(JMIR Aging 2025;8:e66104) doi:10.2196/66104

KEYWORDS

gerontology; geriatrics; older adults; older people; aging; motoric cognitive risk syndrome; MCR; physio-cognitive decline syndrome; PCDS; cognitive frailty; CF; frailty; discrimination; risk factors; prediction; dementia risk; mortality risk

Introduction

The aging process encompasses various physiological declines across multiple systems. As the global population ages, the prevalence of dementia is rising, expected to reach 131 million individuals by 2050. This increase poses a significant economic challenge, potentially equating to 1.1% of the global gross domestic product by 2030 [1]. The lack of curative treatments for dementia and its substantial public health impact underscore the necessity for early detection to mitigate or delay its onset. Given dementia's heterogeneous and lengthy preclinical phase, early screening and diagnosis in at-risk individuals are vital for disease management and caregiver preparedness [2].

Cognitive and physical declines with age often occur simultaneously, suggesting shared underlying mechanisms [3]. Researchers have developed metrics across molecular, phenotypic, and functional areas [4] to reflect the complex nature of aging accurately. Cognitive deterioration typically

precedes dementia by several years, with evidence indicating that motor decline, especially in walking speed, can precede cognitive decline by over a decade [5,6]. Thus, composite aging metrics, encompassing both physical and cognitive functions, offer a promising method to gauge the functional status of the aging population. A meta-analysis highlighted an increased dementia risk in individuals with both physical frailty and cognitive impairment compared to those with cognitive impairment alone [7]. Therefore, composite aging metrics may serve as focal points for interventions aimed at preventing or delaying disability onset and enhancing the healthy lifespan of the elderly [8,9].

An ideal dementia screening tool for primary care should be brief, easily administered, acceptable to older individuals, and exhibit high sensitivity and specificity. Research on aging metrics that incorporate both cognitive and physical functions is gaining traction in gerontological studies due to their strong predictive power for adverse health outcomes. Various metrics have been introduced, such as cognitive frailty (CF) [10], motoric cognitive risk syndrome (MCR) [11], and physio-cognitive decline syndrome (PCDS) [12]. Despite their conceptual similarities, detailed assessments of their definitions and attributes are scarce, hindering their application in research and clinical settings.

Research has shown an increased incidence of concurrent gait and cognitive impairments in older adults susceptible to dementia [5]; however, the specific clinical traits of those experiencing both declines are not well-defined, and direct comparisons between different aging metrics have not been made. Additionally, the effectiveness of these metrics in identifying at-risk individuals and assessing the risk of adverse health events within the same population is not well understood. This lack of knowledge is clinically important for effectively categorizing older adults and identifying potentially reversible conditions in individuals with concurrent declines, thereby informing targeted strategies to slow dementia progression or reduce mortality rates. Consequently, our study seeks to fill these voids by comparing the risk of future dementia and all-cause mortality across 3 aging metrics and examining the predictive abilities of MCR, PCDS, and CF for adverse events among community-dwelling older adults without dementia at baseline, employing data from a broad population-based cohort.

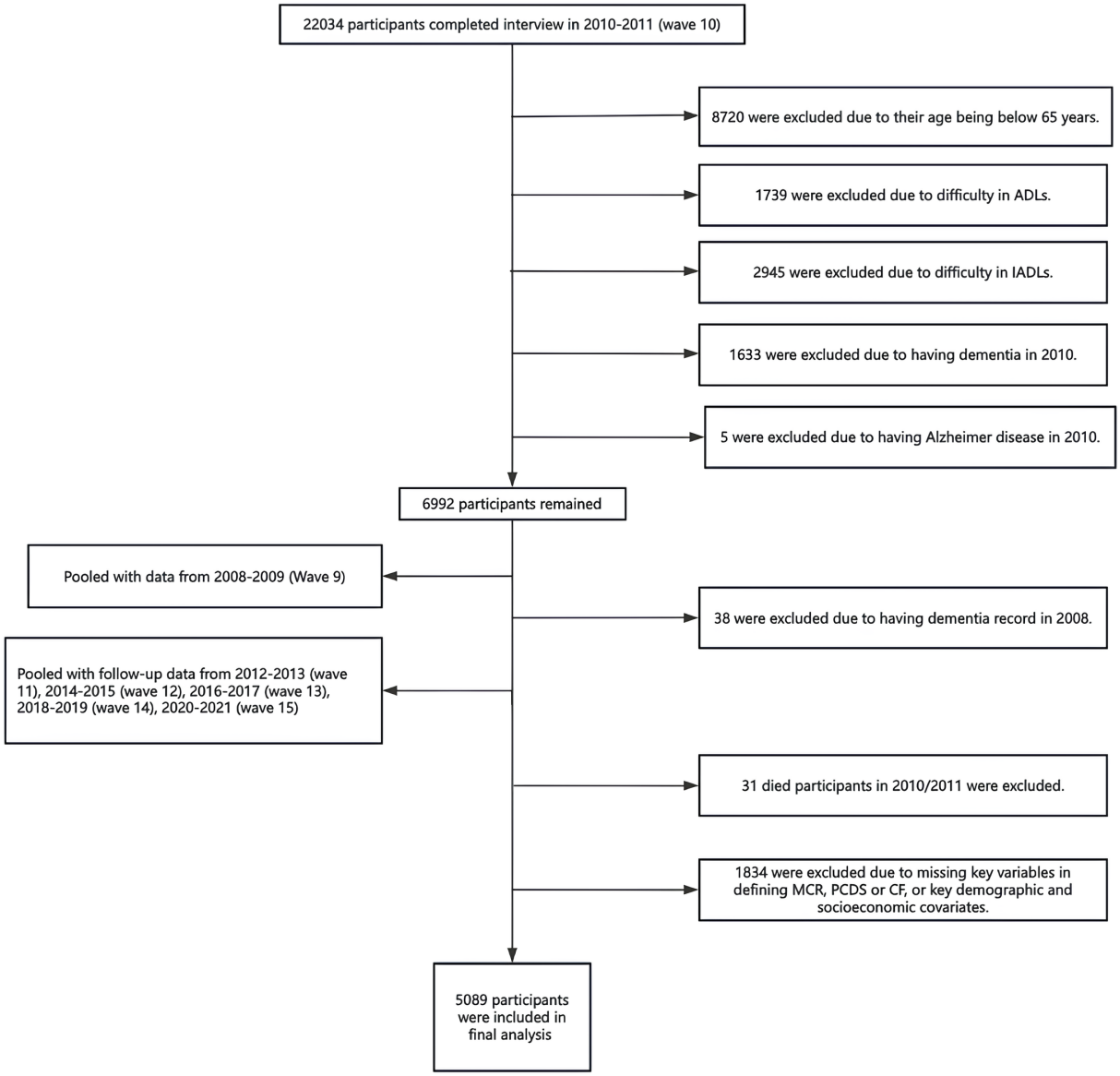
Methods

Sample

This study analyzes data from waves 10 - 15 of the Health and Retirement Study (HRS), the most extensive longitudinal study examining the aging experiences of Americans aged 51 and older. The HRS uses multi-stage probability sampling of U.S. households to obtain a nationally representative sample of adults in this age group [13]. It collects self-reported data on demographics, chronic health conditions, daily activities, disability status, and other health determinants at baseline and every 2 years thereafter. In 2006, the HRS began conducting enhanced face-to-face interviews that included physical performance assessments, biomarker collections, and a leave-behind questionnaire on psychosocial topics. Half of the households were chosen randomly for the enhanced interview in 2006, with the remainder selected in 2008, a process that continues in subsequent waves. Further information on the HRS's recruitment strategies and design is detailed in previous literature [13].

The baseline for this analysis combined data from the 2008 - 2009 (wave 9) and 2010 - 2011 (wave 10) waves, marking the initial occasion respondents were asked about diagnoses of Alzheimer disease or dementia, instead of a "memory-related disease." Mortality data has been available since 2011. A total of 22,034 individuals completed wave 10 and were followed biennially through to 2020 - 2021 (wave 15). This study is a secondary analysis of the de-identified HRS public data, and the original HRS was approved by the University of Michigan Institutional Review Board. All participants signed the informed consent at the time of participation. Our final sample consisted of 2372 individuals who (1) were 65 years or older, (2) had complete baseline data on MCR, PCDS, and CF measures, (3) reported no difficulty with any activities of daily living and instrumental activities of daily living at baseline, (4) did not have Alzheimer disease or dementia at baseline, and (5) were alive in 2010 and 2011. Figure 1 shows the participant flow through each selection stage according to the inclusion criteria.

Figure 1. Study flowchart of participant selection. ADL: activity of daily living; IADL: instrumental activity of daily living; CF: cognitive frailty; MCR: motoric cognitive risk syndrome; PCDS: physio-cognitive decline syndrome.



Measures

Cognitive Function

Biennial cognitive function tests were administered by trained HRS interviewers either in-person or via telephone using the Modified Telephone Interview for Cognitive Status (TICS-m), a global cognition test modeled on the Mini-Mental State Examination. The TICS-m comprises immediate and delayed 10-noun free recall tests (score range: 0 - 10 for each), a serial 7 subtraction test (score range: 0 - 5), and a test of counting backward from 20 (score range: 0 - 2). Higher scores denote better cognitive performance. During each assessment, HRS participants were categorized into normal cognition, mild cognitive impairment (MCI), or dementia based on established thresholds and comprehensive evaluations, including expert clinician adjudication from the Aging, Demographics, and Memory Study, a dementia sub-study within the HRS framework. Cognitive status was categorized into 3 distinct groups based on continuous scores [14], where scores from 12 to 27 indicated no impairment; scores between 7 and 11 signified cognitive impairment without dementia or MCI; and scores from 0 to 6 suggested dementia [15].

Motoric Cognitive Risk Syndrome

MCR was defined by subjective cognitive complaints coupled with slow gait in older adults who did not have a mobility disability or dementia [16]. In the HRS, gait speed (measured in meters per second) was determined from the time taken (in seconds) to walk a 2.5-meter course at a normal pace within participants' homes. A slow gait was defined as performance ≥1 SD below the mean for the participant's age and sex, a criterion previously used in HRS to identify MCR [17].

Subjective cognitive complaints were assessed using 2 questions: 1. "How would you rate your memory at the present time? Would you say it is excellent, very good, good, fair, or poor?" and 2. "Compared with the previous interview, would you say your memory is better now, about the same, or worse now than it was then?" Responses of 'fair' or 'poor' to the first question,

or 'worse' to the second, were considered indicative of cognitive complaints.

Cognitive Frailty

CF was defined as the co-occurrence of physical frailty and MCI [18]. The concept of frailty was identified using the 5 criteria outlined by Fried et al [19] in the Cardiovascular Health Study: unintended weight loss, physical inactivity, exhaustion, weakness, and reduced speed. Unintended weight loss is recognized as either a 10% or greater reduction in BMI since the last measurement in 2008 or a current BMI less than 18.5 kg/m^2. Physical activity levels were quantified by averaging the frequencies of activities at 3 levels of intensity—mild, moderate, and vigorous—weighted by the average metabolic equivalent of task (MET) scores for each intensity level: mild (1 - 3 MET), moderate (4 - 6 MET), and vigorous (7 - 10 MET). Participants were deemed physically inactive if their average physical activity fell within the lowest 20%. Exhaustion was determined based on responses to 2 questions from the Center for Epidemiologic Studies Depression scale [20], including, "I could not get going" and "I felt that everything I did was an effort." Muscle strength was assessed through the average of 2 grip strength measurements using a dynamometer on the dominant hand. Weakness was determined by grip strength falling below thresholds adjusted for BMI and gender, as established in the CHS. Reduced speed or slowness was defined as a speed <0.762 m/s for women taller than 159 cm or men taller than 173 cm and as <0.653 m/s for women 159 cm tall or less or men 173 cm tall or less, measured on a 2.5-m course [19,21]. Participants were considered to have missing data for physical measures if they were unable to perform the assessments due to the absence of suitable facilities or equipment or recent surgery. The diagnosis of frailty was based on the number of these criteria fulfilled, with those meeting 3-5 criteria classified as frail.

Physio-Cognitive Decline Syndrome

PCDS was defined as slowness or weakness (using cutoffs from the 2019 consensus update by the Asian Working Group for Sarcopenia), accompanied by cognitive performance that is at least 1.5 standard deviations below the mean for age-, sex-, and education-matched controls across all cognitive domains [12,22]. This assessment is based on comprehensive objective neuropsychological tests.

All-Cause Dementia and Mortality

The diagnosis of dementia was based on physician-diagnosed dementia and TICS scores ranging from 0 to 6. Mortality data were recorded, including the year and month of death, obtained from an exit interview or a spouse or partner's core interview.

Covariates

Covariates included sociodemographic factors, clinical characteristics, and health-related lifestyle behaviors, all of which were assessed at baseline. Sociodemographic characteristics included age (in years), sex (male or female), educational background (primary school or below, high school or equivalent, college and above), and marital status (married vs unmarried). Health-related lifestyle behaviors included excessive drinking, defined as more than 14 drinks per week for men and more than 7 drinks per week for women. Alcohol consumption was calculated by multiplying the number of days per week that alcohol (drink liquor or beer or wine or rice) was consumed by the number of drinks (liang or bottles or mugs) per day. Clinical characteristics included history of hypertension, diabetes, and heart disease. Hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, physical diagnosis, or antihypertensive medication use [23,24]. Diabetes was defined as having a diabetes diagnosis by a physician, being on treatment for diabetes, and having a fasting glucose level greater than or equal to 126 mg/dL and HbA_{1c} >6.5%. The presence of heart disease was determined via a physician's diagnosis obtained through an in-person visit with study personnel via a questionnaire.

Statistical Analysis

The incidence rates of all-cause dementia and mortality were calculated as the number of incident cases divided by the number of person-years of follow-up within the observation year (from 2008 to 2021). Differences between the MCR and non-MCR groups, PCDS and non-PCDS groups, and CF and non-CF groups were assessed using a 2-sided, independent *t*-test and the χ^2 test. To evaluate the impact of MCR, PCDS, and CF on the occurrence of all-cause dementia and mortality, we employed Cox proportional hazards regression analysis. The observation period extended from the index date to the earliest of the following: onset of dementia, death, or the conclusion of the observation period (December 31, 2018). Adjusted hazard ratios (AHRs) for MCR, PCDS, and CF in predicting the onset of dementia and all-cause mortality were calculated, accounting for covariates in an initially unadjusted model. Subsequent adjustments for covariates were made in 2 stages: Model 1 adjusted for age and gender; Model 2 further incorporated socioeconomic (education level and marital status), lifestyle (excessive drinking), and medical conditions (hypertension, diabetes, heart disease, and stroke). To assess and compare the predictive accuracy of all models, discrimination, defined by the model's ability to differentiate between individuals who develop dementia and those who do not, was quantified using Harrell C-statistic, taking survival into account.

Several sensitivity analyses were conducted to verify the stability of our findings. First, to focus on new cases and reduce reverse causation bias, individuals diagnosed with dementia or who died within 2 years of follow-up (sensitivity analysis I) were excluded. Second, to address the competing risk of death for dementia occurrences, Fine and Gray competing risk models were used [25], comparing these to the results from Cox proportional hazards regression models (sensitivity analysis II). The sub-distribution hazard function, defined at time *t*, represents the immediate risk of event *k* among individuals not previously experiencing event *k*. Third, to minimize selection bias, associations within individual samples with complete data on MCR, PCDS, and CF were analyzed separately (sensitivity analysis III). Statistical analyses were performed using 2-tailed tests with a significance level of $P<.05$ and 95% CIs, employing Stata (version 17) for all statistical procedures.

Ethical Considerations

This investigation has been conducted in accordance with the ethical standards of the Declaration of Helsinki, as well as national and international guidelines. The study has been approved by the Institutional Review Board at the University of Michigan (approval number: HUM00061128). All participants were provided with detailed information about the study, including its purpose, procedures, potential risks and benefits, and their rights to withdraw at any time. Written informed consent was obtained from all participants prior to their involvement in the study. To ensure privacy and confidentiality, all data collected were anonymized and deidentified. No identifying information was retained or published. Protective measures were in place to safeguard participant information, including secure storage of data and restricted access to study records. Participants were not compensated for their involvement in this study. The research was conducted on a voluntary basis, and no financial or other incentives were provided.

Results

Baseline Characteristics

The study participants' baseline characteristics are outlined in Table 1. The initial cohort comprised 2367 individuals, featuring prevalence rates for MCR, PCDS, and CF at 5.4% (n=121), 6.3% (n=140), and 1.3% (n=31), respectively. Among these, CF patients were the oldest on average (75.7, SD 6.1 years), with the distribution of men being 52.89% (64/121) in the MCR group, 47.14% (66/140) in the PCDS group, and 58.06% (10/31) in the CF group. During the follow-up, 573 (24.2%) patients died. The proportions of all-cause dementia for MCR, PCDS, and CF were 24.8% (30/121), 33.6% (47/140), and 41.9% (13/31), respectively.

Table . Characteristics of included patients at baseline according to 3 aging metrics.

Variable	Non-MCR[a] (n=2246)	MCR (n=121)	*P* value	Non-PCDS[b] (n=2227)	PCDS (n=140)	*P* value	Non-CF[c] (n=2336)	CF (n=31)	*P* value
Age (years), mean (SD)	73.78 (5.65)	73.29 (5.54)	.17	73.72 (5.62)	74.29 (6.02)	.08	73.73 (5.63)	75.65 (6.11)	.52
Male, n (%)	1086 (48.35)	64 (52.89)	.33	1084 (48.68）	66 (47.14)	.73	1132 (48.46)	18 (58.06)	.29
Educational background, n (%)			.001			.80			<.001
Illiterate	286 (12.73)	27 (22.31)		297 (13.34）	16 (11.43)		301 (12.89）	12 (38.71)	
Primary or above	1286 (57.26)	73 (60.33)		1276 (57.30%)	83 (59.29)		1343 (57.49)	16 (51.61)	
Secondary or above	674 (30.01)	21 (17.36)		654 (29.37)	41 (29.29)		692 (29.62)	3 (9.68)	
Medical history, n (%)									
Hypertension	1398 (62.24)	91 (75.21)	.004	1392 (62.51)	97 (69.29)	.11	1466 (62.76)	23 (74.19)	.19
Diabetes	431 (19.19)	25 (20.66)	.69	410 (18.41)	46 (32.86)	<.001	446 (19.09)	10 (32.26)	.07
Heart disease	605 (26.94)	47 (38.84)	.004	615 (27.62)	37 (26.43)	.76	638 (27.31)	14 (45.16)	.03
Excessive drink	1436 (63.94)	95 (78.51)	.001	1428 (64.12)	103 (73.57)	.02	1509 (64.60)	22 (70.97)	.46
Incident all-cause dementia	311 (13.85)	30 (24.79)	.001	294 (13.20)	47 (33.57)	<.001	328 (14.04)	13 (41.94)	<.001
Mortality	530 (23.60)	43 (35.54)	.003	526 (23.62)	47 (33.57)	.008	556 (23.80）	17 (54.84)	<.001

[a] MCR: motoric cognitive risk syndrome.

[b]PCDS: physio-cognitive decline syndrome.

[c]CF: cognitive frailty.

Relationships of MCR, PCDS, and CF With Incident Dementia and All-Cause Mortality

Overall, there were 341 incident dementia cases during follow-up, for an overall incidence rate of 19.52 (95% CI 17.56-21.71) per 1000 person-years. There were 573 cases that died during follow-up, for an overall incidence rate of 30.14 (95% CI 27.77-32.71) per 1000 person-years (Table 2). The incidence rates of all-cause dementia among MCR, CF, and PCDS patients were 38.04, 108.11, and 55.80 per 1000 person-years, respectively—significantly higher than those observed in the relatively healthy control group. Similarly, the incidence rates of all-cause mortality were 47.99, 100.79, and 44.70 per 1000 person-years for MCR, CF, and PCDS patients, respectively—again, markedly higher than in the healthy controls. Table 2 also demonstrates significant associations of the 3 conditions with increased risks of incident dementia and all-cause mortality in various models (all *P* values <.001). CF had the highest AHR for both outcomes (5.06; 95% CI 2.87-8.92 for dementia, and 3.26; 95% CI 1.99 - 5.33 for mortality). Participants with dual decline experienced a two to threefold increased risk of dementia progression (AHR: 1.90 - 2.22 in the MCR group; 3.21 - 3.35 in the PCDS group) compared with those without dual decline. The AHR for all-cause mortality ranged from 1.60 to 1.84 in the MCR group and from 1.56 to 1.63 in the PCDS group. The trend was consistent across models: AHRs for dementia and mortality increased from the unadjusted model to adjusted model 1, then decreased upon further adjustment for covariates, including socioeconomic status, lifestyle factors, and medical comorbidities. Sensitivity analyses I (Table S1 in Multimedia Appendix 1) and III (Table S2 in Multimedia Appendix 1) consistently showed an elevated risk of dementia and mortality for MCR, CF, and PCDS. However, in sensitivity analysis II, the standardized hazard ratios for MCR and CF were not statistically significant post-adjustment (Table S3 in Multimedia Appendix 1).

Table . Associations between 3 aging metrics, incident all-cause dementia, and all-cause mortality.

	Total sample	MCR[a]		CF[b]		PCDS[c]	
		No	Yes	No	Yes	No	Yes
All-cause dementia							
Events and sample size, n/N	341/2367	311/2246	30/121	328/2336	13/31	294/2227	47/140
Incidence (95% CI)[d]	19.52 (17.56 - 21.71)	18.65 (16.69 - 20.84)	38.04 (26.59 - 54.40)	18.91 (16.97 - 21.07)	108.11 (62.77 - 186.18)	17.68 (15.77 - 19.83)	55.80 (41.93 - 74.27)
Unadjusted HR[e] (95% CI)	—[f]	Ref.[g]	2.03 (1.40 - 2.95)	Ref.	6.23 (3.57 - 10.88)	Ref.	3.21 (2.36 - 4.37)
Model 1[h]: adjusted HR (95% CI)	—	Ref.	2.22 (1.53 - 3.24)	Ref.	6.62 (3.79 - 11.57)	Ref.	3.29 (2.41 - 4.48)
Model 2[i]: adjusted HR (95% CI)	—	Ref.	1.90 (1.30 - 2.78)	Ref.	5.06 (2.87 - 8.92)	Ref.	3.35 (2.44 - 4.58)
All-cause mortality							
Events and sample size, n/N	573/2367	530/2246	43/121	556/2336	17/31	526/2227	47/140
Incidence (95% CI)	30.14 (27.77 - 32.71)	29.26 (26.87 - 31.86)	47.99 (35.59 - 64.71)	29.51 (27.16 - 32.07)	100.79 (62.66 - 162.13)	29.29 (26.89 - 31.90)	44.70 (33.58 - 59.49)
Unadjusted HR (95% CI)	—	Ref.	1.68 (1.23 - 2.29)	Ref.	3.93 (2.42 - 6.38)	Ref.	1.56 (1.16 - 2.11)
Model 1: adjusted HR (95% CI)	—	Ref.	1.84 (1.35 - 2.51)	Ref.	4.04 (2.49 - 6.57)	Ref.	1.63 (1.21 - 2.19)
Model 2: adjusted HR (95% CI)	—	Ref.	1.60 (1.17 - 2.19)	Ref.	3.26 (1.99 - 5.33)	Ref.	1.58 (1.17 - 2.13)

[a]MCR: motoric cognitive risk syndrome.

[b]CF: cognitive frailty.

[c]PCDS: physio-cognitive decline syndrome.

[d]Incidence rates = events per 1000 person-years.

[e]HR: hazard ratio.

[f]Not applicable.

[g]Ref: reference.

[h]Model 1 adjusted for age and gender.

[i]Model 2 further adjusted for educational background, marital status, excessive drinking, hypertension, diabetes, and heart disease.

Discriminations of MCR, PCDS, and CF for All Outcomes

The discrimination abilities of MCR, PCDS, and CF regarding all-cause dementia and mortality are detailed in Table 3. The C-index for PCDS in identifying all-cause dementia was 0.732 (95% CI 0.703-0.760), outperforming the C-indices for MCR and CF (PCDS vs MCR: 0.012, 95% CI −0.001 to 0.025; *P*=.08; CF vs MCR: 0.005, 95% CI −0.004 to 0.013; *P*=.35). In contrast, MCR's C-index for identifying all-cause mortality was the highest among the 3, at 0.727 (95% CI 0.706-0.748). Similar to the case with incident dementia, the differences in discrimination ability among MCR, PCDS, and CF for all-cause mortality were not statistically significant.

Table . Harrell C-index for Cox regression models predicting incident all-cause dementia, and all-cause mortality.

	C-index[a]	95% CI	*P* value	Difference	*P* value
Incident all-cause dementia					
MCR[b]	0.7194	0.6915 to 0.7472	<.001	Reference	
PCDS[c]	0.7315	0.7028 to 0.7602	<.001	0.0121 (−0.0009 to 0.0252)	.08
CF[d]	0.7239	0.6959 to 0.7520	<.001	0.0046 (−0.0038 to 0.0129)	.35
All-cause mortality					
MCR	0.727	0.7061 to 0.7479	<.001	Reference	
PCDS	0.7259	0.7048 to 0.7469	<.001	−0.0012 (−0.0054 to 0.0030)	.59
CF	0.7254	0.7044 to 0.7465	<.001	−0.0016 (−0.0054 to 0.0022)	.41

[a]All indexes were estimated in models adjusted for age, gender, educational background, marital status, excessive drinking, hypertension, diabetes, and heart disease

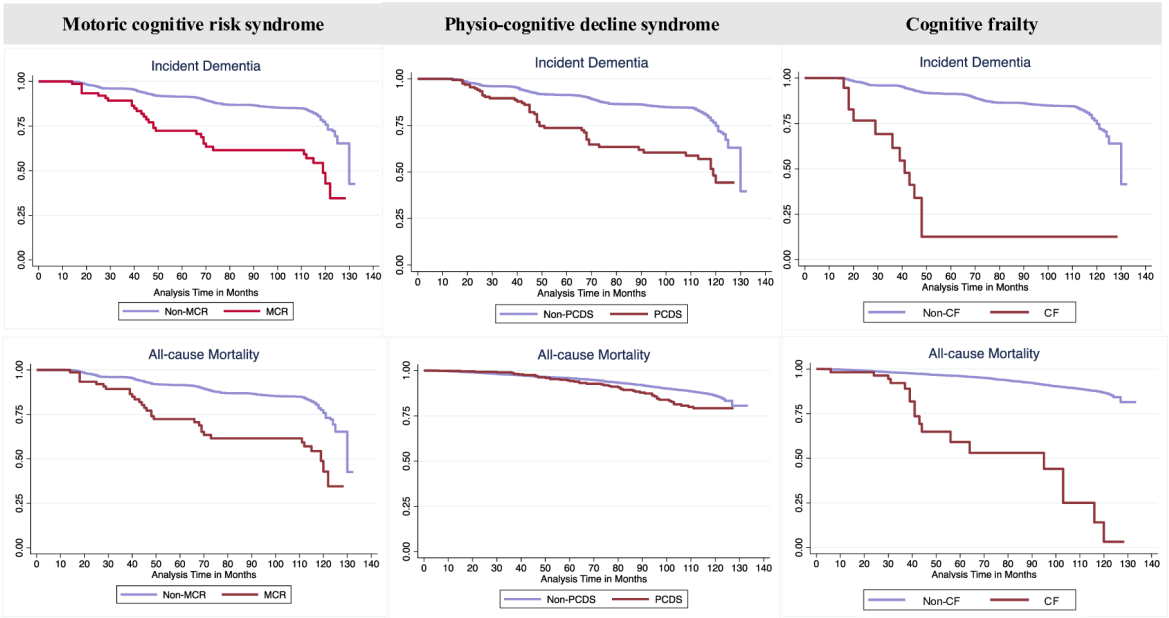
[b]MCR: motoric cognitive risk syndrome.

[c]PCDS: physio-cognitive decline syndrome.

[d]CF: cognitive frailty.

The Kaplan-Meier curve, depicted in Figure 2, shows the duration to incident dementia or all-cause mortality, stratified by MCR, PCDS, or CF, and healthy controls, with adjustments made for all covariates. Both curves demonstrate a decline over the follow-up period, with a notably pronounced decrease observed among patients (all *P* values for log rank <.001).

Figure 2. Kaplan-Meier survival curve depicting the proportion of participants remaining dementia-free or surviving during follow-up, comparing motoric cognitive risk syndrome (MCR), physio-cognitive decline syndrome (PCDS), and cognitive frailty (CF) patients with their respective healthy control groups.



Discussion

Principal Findings

Analyzing data from a large-scale cohort in the United States, this study is the first to illustrate that MCR, PCDS, and CF have significant associations with incident dementia and all-cause mortality within the same population, despite the inherent heterogeneity of these metrics. Among them, CF emerged as the highest-risk group for both dementia and mortality, with markedly elevated incidence rates and the highest AHRs (5.06 for dementia and 3.26 for mortality), identifying CF as the most vulnerable subgroup. These findings offer valuable insights into the future health trajectories of frail individuals with cognitive impairments and highlight the need for targeted interventions and early health education to mitigate dementia and mortality risks in later life at lower costs. While PCDS showed the strongest predictive ability for dementia and MCR for mortality, the differences in predictive accuracy were not statistically significant. Implementing appropriate management strategies can help alleviate the health care burden for individuals with varying cognitive and physical conditions.

To forestall unhealthy aging, health care systems need to identify individuals at risk of functional declines that are still preventable or reversible. Our results are in line with prior studies, indicating that MCR syndrome is linked with a 90% increased risk of incident dementia and a 60% higher risk of mortality, even after adjustments for demographics, SES, and cardiovascular comorbidities (hypertension, diabetes, and heart disease). Consistent with research from Chung et al, the prevalence of hypertension and heart disease was significantly higher in the MCR group, highlighting the importance of regular screening to manage cardiovascular risk factors and delay dementia onset [26,27]. These observations further emphasize the potential neurodegenerative nature of MCR pathologic changes [28].

CF has been associated with an increased risk of dementia and mortality, as demonstrated by a meta-analysis revealing that older adults with CF face higher mortality and dementia risks than their healthier counterparts [29]. However, the absence of a standardized assessment tool for CF hampers its use in widespread epidemiological studies. Using the frailty phenotype alongside the MMSE, although common, proves to be time-intensive and presents challenges for deployment in community and hospital settings [30]. Furthermore, the lower prevalence of CF compared to other measures (1.31% for CF vs 5.91% for PCDS vs 5.11% for MCR) in the community-dwelling older adult cohort indicates difficulties in identifying a substantial number of at-risk individuals for interventions in similar groups.

In contrast to CF, PCDS, with its precise definition encompassing physical decline and cognitive impairment, provides a more targeted approach. By focusing on the mobility aspects of frailty linked to worse cognitive outcomes and increased mortality risk, researchers can more accurately identify the target population, concentrating on specific causal mechanisms [31,32]. PCDS serves as an advantageous focus for multidomain interventions, integrating physical activity, cognitive training, nutritional counseling, and disease management education to improve the condition of vulnerable but potentially reversible older individuals [33]. A recent Singapore study that used a 12-week dual-task exercise program to examine the potential reversibility of CF, MCR, and PCDS also found clinical improvements in PCDS, but the longer-term effects remain uncertain [34]. Nonetheless, Lee et al [35] demonstrated that PCDS was related to a 6-year, but not 3-year, incidence of dementia, even when using inverse probability weighting analysis to account for bias from missing data. This finding implies a latency period for the development of dementia, highlighting this interval as potentially crucial for preventive interventions.

While PCDS demonstrated the best predictive performance for incident dementia, MCR emerged as a superior predictor for all-cause mortality. The choice between these 2 metrics may depend on practical considerations, with MCR offering several unique advantages for clinical use. MCR is simpler and more efficient to assess, unaffected by education levels or learning effects, which enhances its credibility and reliability as a screening tool. In contrast, the comprehensive cognitive assessments required for PCDS diagnosis can be labor-intensive and require specific skills, which may limit their use in community or primary care settings. Given MCR's higher predictive accuracy for all-cause mortality and its ease of implementation, it is well-suited for use in routine health care assessments, particularly in primary health care settings. Additionally, research suggests that while handgrip strength declines earlier in aging, walking speed in later life is a stronger predictor of mortality [36]. This further supports MCR's higher predictive accuracy for all-cause mortality.

Despite representing a prodromal phenotype of accelerated aging, MCR remains an important target for intervention to prevent poor outcomes in older adults. MCR offers incremental validity in predicting dementia beyond what is provided by MCI subtypes [16] and individual components such as subjective cognitive complaints or slow gait [16,37]. Our findings suggest that MCR could serve as a simpler, more efficient tool to identify individuals at higher risk for dementia, especially given its independence from educational background or learning effects. The motor components of MCR, such as the time-up-and-go test or one-leg-standing test, are reliable and valid methods for quantifying motor function, making MCR an easily implementable screening tool in various health care settings. However, one study has shown that the MCR subtype defined by the 5-times-sit-to-stand test, which also includes a balance component, was less effective in predicting cognitive decline compared to MCR defined by slow gait [38]. Further research is needed to determine how different gait and balance assessments influence the identification of individuals at risk for dementia, and whether integrating MCR into routine geriatric assessments could enhance early intervention strategies in aging populations.

Early and appropriate intervention might postpone the onset of dementia or reduce its risk, particularly given the lengthy incubation period associated with the development of dementia. Our results could aid clinicians in the early implementation of screening and prevention strategies and inform government decisions on community health prevention. The strength of this

study lies in its extensive, well-delineated cohort, featuring longitudinal assessments, uniform measures, and verified outcomes. To our knowledge, this is the first study to assess the predictive accuracy of MCR, PCDS, and CF regarding dementia incidence and all-cause mortality within the same cohort. The reliability and consistency of our findings, supported by sensitivity analyses, bolster the study's validity. Nonetheless, several limitations should be acknowledged. First, the study lacked objective neuropsychological testing [39] and did not account for some potential confounding factors, such as *APOE* genotype or imaging biomarkers [40,41]. However, our sensitivity analysis—excluding participants diagnosed with dementia or those who died within 2 years of the index date—strengthened the robustness of our conclusions. Second, while we used data from a large longitudinal cohort, the final sample size of 2372 individuals may limit the representativeness of our findings. Additionally, the prevalence rates of MCR (5.4%), PCDS (6.3%), and CF (1.3%) in our sample may be higher than those in the general older population, introducing potential selection bias. To address this, sensitivity analyses were conducted within individual subgroups with complete data on MCR, PCDS, and CF. These analyses included larger sample sizes and showed lower prevalence rates for the 3 conditions, aligning more closely with those observed in the general population. The robustness of these results further confirmed our conclusions, enhancing their validity and generalizability while mitigating potential selection bias. Future research could benefit from larger, more diverse, and representative samples to enhance generalizability. Third, the potential influence of mortality on dementia risk assessment must be considered. Individuals who died during follow-up could bias dementia outcome evaluations. To address this competing risk, we employed Fine and Gray competing risk models and compared the results with those from Cox proportional hazards regression models as part of our sensitivity analysis. This approach ensured that our findings remained robust and accounted for the impact of mortality on dementia risk associations. Finally, variations in syndrome definitions and differences in the prevalence of the 3 aging metrics within the population could affect their predictive power for adverse outcomes. This variability may limit the generalizability of our results to other populations. To mitigate this, sensitivity analyses within subgroups with complete data on MCR, PCDS, and CF were conducted, reinforcing the reliability of the associations between these aging metrics and adverse health outcomes.

Conclusions

The integration of cognitive and physical functions into aging metrics consistently indicates risks for incident dementia and all-cause mortality, despite their significant differences, underscoring their utility for effective risk stratification in research and clinical settings. Patients with CF represent the most vulnerable subgroup, highlighting the need for prioritized preventive strategies and interventions. Meanwhile, MCR emerged as a particularly efficient and accurate screening tool for both dementia and mortality. These findings emphasize the importance of implementing targeted prevention and intervention programs based on these metrics to enhance quality of life and reduce premature mortality among aging populations. Further research on the longitudinal dynamics of these aging metrics in relation to dementia, mortality, and other outcomes is essential for a deeper understanding of their long-term impact.

Acknowledgments

This study was supported by the Natural Science Foundation of Guangdong Province of China (2023A1515011259).

Data Availability

The datasets analyzed during this study are available from the Health and Retirement Study [42].

Authors' Contributions

AYB and WHX contributed to conception and design. AYB and SH contributed to writing the original draft and reviewing and editing the manuscript. AYB, YJ, and SH were involved in data curation, and data analysis. YJ and ZYL contributed to reviewing and editing. AYB and WHX contributed to the methodology, project administration, supervision, and verified the underlying data. All authors gave final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File, 26 KB - aging_v8i1e66104_app1.docx]

References

1. Zeisel J, Bennett K, Fleming R. World Alzheimer report 2020: design, dignity, dementia: dementia-related design and the built environment. 2020.
2. Borson S, Frank L, Bayley PJ, et al. Improving dementia care: the role of screening and detection of cognitive impairment. Alzheimers Dement 2013 Mar;9(2):151-159. [doi: 10.1016/j.jalz.2012.08.008] [Medline: 23375564]

3. Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment--a review of the evidence and causal mechanisms. Ageing Res Rev 2013 Sep;12(4):840-851. [doi: 10.1016/j.arr.2013.06.004] [Medline: 23831959]
4. Ferrucci L, Levine ME, Kuo PL, Simonsick EM. Time and the metrics of aging. Circ Res 2018 Sep 14;123(7):740-744. [doi: 10.1161/CIRCRESAHA.118.312816] [Medline: 30355074]
5. Montero-Odasso MM, Barnes B, Speechley M, et al. Disentangling cognitive-frailty: results from the gait and brain study. J Gerontol A Biol Sci Med Sci 2016 Nov;71(11):1476-1482. [doi: 10.1093/gerona/glw044] [Medline: 26984391]
6. Dumurgier J, Artaud F, Touraine C, et al. Gait speed and decline in gait speed as predictors of incident dementia. J Gerontol A Biol Sci Med Sci 2016;72(5). [doi: 10.1093/gerona/glw110]
7. Zheng L, Li G, Gao D, et al. Cognitive frailty as a predictor of dementia among older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2020;87:103997. [doi: 10.1016/j.archger.2019.103997] [Medline: 31846833]
8. Cano A. Cognitive frailty, a new target for healthy ageing. Maturitas 2015 Oct;82(2):139-140. [doi: 10.1016/j.maturitas.2015.07.026] [Medline: 26277253]
9. Ruan Q, Yu Z, Chen M, Bao Z, Li J, He W. Cognitive frailty, a novel target for the prevention of elderly dependency. Ageing Res Rev 2015 Mar;20:1-10. [doi: 10.1016/j.arr.2014.12.004] [Medline: 25555677]
10. Kelaiditi E, Cesari M, Canevelli M, et al. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) International Consensus Group. J Nutr Health Aging 2013 Nov;17(9):726-734. [doi: 10.1007/s12603-013-0367-2]
11. Verghese J, Annweiler C, Ayers E, et al. Motoric cognitive risk syndrome: multicountry prevalence and dementia risk. Neurology (ECronicon) 2014 Aug 19;83(8):718-726. [doi: 10.1212/WNL.0000000000000717] [Medline: 25031288]
12. Chung CP, Lee WJ, Peng LN, et al. Physio-cognitive decline syndrome as the phenotype and treatment target of unhealthy aging. J Nutr Health Aging 2021;25(10):1179-1189. [doi: 10.1007/s12603-021-1693-4] [Medline: 34866145]
13. Heeringa SG, Connor JH. Technical Description of the Health and Retirement Survey Sample Design: University of Michigan; 1995.
14. Langa KM, Plassman BL, Wallace RB, et al. The Aging, Demographics, and Memory Study: study design and methods. Neuroepidemiology 2005;25(4):181-191. [doi: 10.1159/000087448] [Medline: 16103729]
15. Crimmins EM, Kim JK, Langa KM, Weir DR. Assessment of cognition using surveys and neuropsychological assessment: the Health and Retirement Study and the Aging, Demographics, and Memory Study. J Gerontol B Psychol Sci Soc Sci 2011 Jul;66(Suppl 1):i162-i171. [doi: 10.1093/geronb/gbr048] [Medline: 21743047]
16. Verghese J, Wang C, Lipton RB, Holtzer R. Motoric cognitive risk syndrome and the risk of dementia. J Gerontol A Biol Sci Med Sci 2013 Apr;68(4):412-418. [doi: 10.1093/gerona/gls191] [Medline: 22987797]
17. Ayers E, Verghese J. Motoric cognitive risk syndrome and risk of mortality in older adults. Alzheim Dementia 2016 May;12(5):556-564. [doi: 10.1016/j.jalz.2015.08.167]
18. Gaspar PM, Campos-Magdaleno M, Pereiro AX, Facal D, Juncos-Rabadán O. Cognitive reserve and mental health in cognitive frailty phenotypes: Insights from a study with a Portuguese sample. Front Psychol 2022;13:968343. [doi: 10.3389/fpsyg.2022.968343] [Medline: 36110284]
19. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
20. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
21. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. J Gerontol A Biol Sci Med Sci 2006 Mar;61(3):262-266. [doi: 10.1093/gerona/61.3.262] [Medline: 16567375]
22. Chen LK, Arai H. Physio-cognitive decline as the accelerated aging phenotype. Arch Gerontol Geriatr 2020;88:104051. [doi: 10.1016/j.archger.2020.104051] [Medline: 32278485]
23. Chen S, Sudharsanan N, Huang F, Liu Y, Geldsetzer P, Bärnighausen T. Impact of community based screening for hypertension on blood pressure after two years: regression discontinuity analysis in a national cohort of older adults in China. Br Med J 2019 Jul 11;366:l4064. [doi: 10.1136/bmj.l4064] [Medline: 31296584]
24. Geldsetzer P, Manne-Goehler J, Marcus ME, et al. The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from 1·1 million adults. Lancet 2019 Aug 24;394(10199):652-662. [doi: 10.1016/S0140-6736(19)30955-9] [Medline: 31327566]
25. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999 Jun;94(446):496-509. [doi: 10.1080/01621459.1999.10474144]
26. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet 2015 Jun;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5]
27. Crous-Bou M, Minguillón C, Gramunt N, Molinuevo JL. Alzheimer's disease prevention: from risk factors to early intervention. Alz Res Therapy 2017 Dec;9(1):71. [doi: 10.1186/s13195-017-0297-z]
28. Sekhon H, Allali G, Launay CP, et al. Motoric cognitive risk syndrome, incident cognitive impairment and morphological brain abnormalities: systematic review and meta-analysis. Maturitas 2019 May;123:45-54. [doi: 10.1016/j.maturitas.2019.02.006] [Medline: 31027677]

29. Bu Z, Huang A, Xue M, Li Q, Bai Y, Xu G. Cognitive frailty as a predictor of adverse outcomes among older adults: a systematic review and meta-analysis. Brain Behav 2021 Jan;11(1):e01926. [doi: 10.1002/brb3.1926] [Medline: 33159430]
30. Zhang XM, Wu XJ, Cao J, Jiao J, Chen W. Association between cognitive frailty and adverse outcomes among older adults: a meta-analysis. J Nutr Health Aging 2022;26(9):817-825. [doi: 10.1007/s12603-022-1833-5] [Medline: 36156673]
31. Wu YH, Liu LK, Chen WT, et al. Cognitive function in individuals with physical frailty but without dementia or cognitive complaints: results from the I-Lan Longitudinal Aging Study. J Am Med Dir Assoc 2015 Oct 1;16(10):899. [doi: 10.1016/j.jamda.2015.07.013] [Medline: 26321467]
32. Liu LK, Chen CH, Lee WJ, et al. Cognitive frailty and its association with all-cause mortality among community-dwelling older adults in Taiwan: results from I-Lan Longitudinal Aging Study. Rejuvenation Res 2018 Dec;21(6):510-517. [doi: 10.1089/rej.2017.2038] [Medline: 29644921]
33. Chen LK, Hwang AC, Lee WJ, et al. Efficacy of multidomain interventions to improve physical frailty, depression and cognition: data from cluster-randomized controlled trials. J Cachexia Sarcopenia Muscle 2020 Jun;11(3):650-662. [doi: 10.1002/jcsm.12534] [Medline: 32134208]
34. Merchant RA, Chan YH, Hui RJY, et al. Motoric cognitive risk syndrome, physio-cognitive decline syndrome, cognitive frailty and reversibility with dual-task exercise. Exp Gerontol 2021 Jul 15;150:111362. [doi: 10.1016/j.exger.2021.111362] [Medline: 33887381]
35. Lee WJ, Peng LN, Lin MH, et al. Six-year transition of physio-cognitive decline syndrome: results from I-Lan Longitudinal Aging Study. Arch Gerontol Geriatr 2022;102:104743. [doi: 10.1016/j.archger.2022.104743] [Medline: 35687948]
36. Zaccardi F, Franks PW, Dudbridge F, Davies MJ, Khunti K, Yates T. Mortality risk comparing walking pace to handgrip strength and a healthy lifestyle: a UK Biobank study. Eur J Prev Cardiol 2021 Jul 10;28(7):704-712. [doi: 10.1177/2047487319885041] [Medline: 34247229]
37. Verghese J, Lipton RB, Hall CB, Kuslansky G, Katz MJ, Buschke H. Abnormality of gait as a predictor of non-Alzheimer's dementia. N Engl J Med 2002 Nov 28;347(22):1761-1768. [doi: 10.1056/NEJMoa020441] [Medline: 12456852]
38. Sekhon H, Launay CP, Chabot J, Allali G, Beauchet O. Motoric cognitive risk syndrome: could it be defined through increased five-times-sit-to-stand test time, rather than slow walking speed? Front Aging Neurosci 2018;10:434. [doi: 10.3389/fnagi.2018.00434] [Medline: 30766485]
39. Alexander LK, Lopes B, Ricchetti-Masterson K, Yeatts KB. Sources of systematic error or bias: Information bias. ERIC Notebook 2015;2(14):1-5.
40. Baumgart M, Snyder HM, Carrillo MC, Fazio S, Kim H, Johns H. Summary of the evidence on modifiable risk factors for cognitive decline and dementia: a population-based perspective. Alzheimers Dement 2015 Jun;11(6):718-726. [doi: 10.1016/j.jalz.2015.05.016] [Medline: 26045020]
41. Campbell NL, Unverzagt F, LaMantia MA, Khan BA, Boustani MA. Risk factors for the progression of mild cognitive impairment to dementia. Clin Geriatr Med 2013 Nov;29(4):873-893. [doi: 10.1016/j.cger.2013.07.009] [Medline: 24094301]
42. HRS data portal. Health and Retirement Study. URL: https://hrsdata.isr.umich.edu/ [accessed 2025-01-28]

Abbreviations

AHR: adjusted hazard ratio
CF: cognitive frailty
HRS: Health and Retirement Study
MCI: mild cognitive impairment
MCR: motoric cognitive risk syndrome
MET: metabolic equivalent of task
OR: odds ratio
PCDS: physio-cognitive decline syndrome
TICS-m: Modified Telephone Interview for Cognitive Status

Edited by Y Du; submitted 03.09.24; peer-reviewed by L Qin, M Colombo, S Wang; revised version received 27.11.24; accepted 05.12.24; published 30.01.25.

Please cite as:
Bai A, He S, Jiang Y, Xu W, Lin Z
Comparison of 3 Aging Metrics in Dual Declines to Capture All-Cause Dementia and Mortality Risk: Cohort Study
JMIR Aging 2025;8:e66104
URL: https://aging.jmir.org/2025/1/e66104
doi:10.2196/66104

Original Paper

Building Consensus on the Relevant Criteria to Screen for Depressive Symptoms Among Near-Centenarians and Centenarians: Modified e-Delphi Study

Carla Gomes da Rocha[1,2,3], MScN; Armin von Gunten[2], MPhil, MD; Pierre Vandel[2], PhD; Daniela S Jopp[4,5], PhD; Olga Ribeiro[6], PhD; Henk Verloo[1], PhD

[1]School of Health Sciences, University of Applied Sciences and Arts Western Switzerland (HES-SO), Sion, Switzerland

[2]Service of Old Age Psychiatry, Lausanne University Hospital and University of Lausanne, Prilly-Lausanne, Switzerland

[3]Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal

[4]Institute of Psychology, University of Lausanne, Lausanne, Switzerland

[5]Swiss National Centre of Competence in Research LIVES, University of Lausanne, Lausanne, Switzerland

[6]Nursing School of Porto (ESEP) and RISE-Health, Porto, Portugal

Corresponding Author:
Carla Gomes da Rocha, MScN
School of Health Sciences
University of Applied Sciences and Arts Western Switzerland (HES-SO)
Chemin de l'Agasse 5
Sion, 1950
Switzerland
Phone: 41 58 606 84 73
Email: carla.gomesdarocha@hevs.ch

Abstract

Background: The number of centenarians worldwide is expected to increase dramatically, reaching 3.4 million by 2050 and >25 million by 2100. Despite these projections, depression remains a prevalent yet underdiagnosed and undertreated condition among this population that carries significant health risks.

Objective: This study aimed to identify and achieve consensus on the most representative signs and symptoms of depression in near-centenarians and centenarians (aged ≥95 years) through an e-Delphi study with an international and interdisciplinary panel of experts. Ultimately, the outcomes of this study might help create a screening instrument that is specifically designed for this unique population.

Methods: A modified e-Delphi study was carried out to achieve expert consensus on depressive symptoms in near-centenarians and centenarians. A panel of 28 international experts was recruited. Consensus was defined as 70% agreement on the relevance of each item. Data were collected through a web-based questionnaire over 3 rounds. Experts rated 104 items that were divided into 24 dimensions and 80 criteria to identify the most representative signs and symptoms of depression in this age group.

Results: The panel consisted of experts from various countries, including physicians with experience in old age psychiatry or geriatrics as well as nurses and psychologists. The response rate remained consistent over the rounds (20/28, 71% to 21/28, 75%). In total, 4 new dimensions and 8 new criteria were proposed by the experts, and consensus was reached on 86% (24/28) of the dimensions and 80% (70/88) of the criteria. The most consensual potentially relevant dimensions were *lack of hope* (21/21, 100%), *loss of interest* (27/28, 96%), *lack of reactivity to pleasant events* (27/28, 96%), *depressed mood* (26/28, 93%), and *previous episodes of depression or diagnosed depression* (19/21, 90%). In addition, the most consensual potentially relevant criteria were *despondency, gloom, and despair* (25/25, 100%); *depressed* (27/27, 100%); *lack of reactivity to pleasant events or circumstances* (28/28, 100%); *suicidal ideation* (28/28, 100%); *suicide attempt(s)* (28/28, 100%); *ruminations* (27/28, 96%); *recurrent thoughts of death or suicide* (27/28, 96%); *feelings of worthlessness* (25/26, 96%); *critical life events* (20/21, 95%); *anhedonia* (20/21, 95%); *loss of interest in activities* (26/28, 93%); *loss of pleasure in activities* (26/28, 93%); and *sadness* (24/26, 92%). Moreover, when assessing depression in very old age, the duration, number, frequency, and severity of signs and symptoms should also be considered, as evidenced by the high expert agreement.

Conclusions: The classification of most elements as *relevant* highlights the importance of a multidimensional approach for optimal depression screening among individuals of very old age. This study offers a first step toward improving depression assessment in near-centenarians and centenarians. The development of a more adapted screening tool could improve early detection and intervention, enhancing the quality of mental health care for this population.

(JMIR Aging 2025;8:e64352) doi:10.2196/64352

KEYWORDS

centenarians; near-centenarians; depressive symptoms; depression diagnosis; screening; assessment; e-Delphi technique; web-based survey

Introduction

Background

The centenarian population is growing significantly and is projected to increase sharply in the coming decades. As of 2024, there were approximately 722,000 centenarians worldwide, with Japan, the United States, and China having the highest numbers [1]. This figure is expected to rise to 3.4 million by 2050 and exceed 25 million by 2100 [2,3]. In Europe, notable examples include France, Italy, and Greece, each with >20 centenarians per 100,000 inhabitants registered already in 2011. Portugal also stands out with a density of 14.4 centenarians per 100,000 inhabitants that same year, ranking it among the top 15 in Europe [4]. In Switzerland, the centenarian population more than doubled from 787 in 2000 to 1948 in 2022, reaching a noteworthy prevalence of 22.1 per 100,000 inhabitants [5].

Amid extreme longevity, individuals' health experiences vary significantly, from robust health to substantial declines [6-8]. Depression, a prevalent mood disorder among older adults [9], emerges as a major public health concern [10]. The *International Classification of Diseases, 11th Revision* characterizes a depressive episode as "a period of depressed mood or diminished interest in activities occurring most of the day, nearly every day during a period lasting at least two weeks accompanied by other symptoms such as difficulty concentrating, feelings of worthlessness or excessive or inappropriate guilt, hopelessness, recurrent thoughts of death or suicide, changes in appetite or sleep, psychomotor agitation or retardation, and reduced energy or fatigue" [6]. However, this definition broadly encompasses the general population and may not fully address the possible difference and heterogeneity of depressive symptomatology in older adults, potentially leading to diagnostic imprecision.

Accurate diagnosis is crucial as depression should not be misconstrued as an inherent aspect of aging even in very advanced age [11,12]. A systematic literature review reported substantial variations in depression prevalence among near-centenarians and centenarians across countries, such as 12.8% in Italy, 13.5% in Australia, 20% in the United States, and 29% in Mexico [13]. These differences may reflect cultural, health care, and socioeconomic influences as well as methodological variations in study conduct [13,14]. Interestingly, comparisons between centenarian cohorts and younger groups have shown inconsistent findings, with half of the studies reporting higher rates of depressive symptoms among centenarians [13]. This highlights the diversity of aging experiences and suggests that mental health may not necessarily worsen with age. Particularly noteworthy are the findings from Sweden, where near-centenarians and centenarians exhibited almost double the depression rates of older adults aged 85 years at 32.3% versus 16.8%, respectively. These statistics highlight the need not only for careful mental health evaluations among individuals of very old age but also for prompt discussion on tailored approaches to support mental well-being in this population [13,15].

Depression can lead to severe physical, cognitive, and psychological effects. Physically, it can cause malnutrition, falls, delirium, functional decline, and increased mortality [13,16-20]. Cognitively, depression has been connected to cognitive decline and dementia [21-25]. Psychologically, it has been associated with anxiety and suicidal ideation [26,27]. Furthermore, depression places a significant burden on caregivers and health care systems [16-18,28,29], increasing costs and resource demands. These consequences underscore the critical need for early detection and intervention. Therefore, proactive measures such as regular mental health screenings are essential to prevent the severe impact of depression on individuals' quality of life and alleviate its wider societal effects [11]; however, these are not part of the current practice in most countries worldwide.

Addressing this gap in care is difficult as depression in older adults frequently remains underdiagnosed and undertreated, in part because its manifestations are often misinterpreted as natural aspects of aging [18,19,30,31]. Studies have estimated that approximately half of depression cases are not identified by frontline health care providers [16,18], and of those detected, approximately half do not receive adequate treatment [16]. The challenge in detecting late-life depression (LLD) stems from its sometimes subtle, hidden, or atypical signs and symptoms [30,32]. Physical manifestations such as weight loss, psychomotor retardation, and exhaustion, as well as emotional manifestations such as loss of interest in activities, heightened anxiety, or irritability, might overshadow the typical sadness often associated with depression [9,18]. This complexity requires greater clinical vigilance. Moreover, the possible overlap of symptoms between depression and dementia adds another layer of diagnostic difficulty as both conditions can present with symptoms such as memory impairment, psychomotor retardation, and reduced motivation [9]. Beyond dementia, several other medical conditions can mimic or exacerbate depressive symptoms. Chronic pain, thyroid dysfunction, Parkinson disease, stroke, and medication side effects—particularly in polymedicated individuals—can contribute to *depression-like* symptoms, including fatigue, apathy, functional dependence, memory impairment, and

disturbances in mood and sleep [33-38]. These overlapping symptoms underscore the need for a comprehensive assessment to distinguish LLD from other conditions, aiming for an accurate differential diagnosis. The wide range of signs and symptoms associated with LLD can extend beyond the conventional diagnostic criteria for depressive disorders. Symptom clusters not fully meeting the diagnostic criteria for depression may still carry significant clinical importance due to their link to reduced quality of life and increased disability [9,18].

Objectives

Given the prevalence of depression among individuals of very old age, better detection of early signs and symptoms is crucial. However, despite the availability of various screening tools, to the best of our knowledge, none have been specifically validated for near-centenarians and centenarians (aged ≥95 years). Therefore, our study aimed to compile a comprehensive range of potentially relevant depressive features that have been documented in scientific literature. These features were then submitted to an international panel of experts to achieve consensus on the most representative signs and symptoms of depression in individuals of very old age. The outcomes of this approach are 2-fold. First, by systematically examining and integrating diverse diagnostic criteria and screening tools, this study aimed to contribute to the ongoing debate about the complexities and challenges of effectively diagnosing depression in individuals of very old age. Second, achieving expert consensus is central to developing a new screening instrument specifically tailored to this unique population. This effort might be an important first step toward early detection and effective intervention, which might lead to better outcomes and enhanced general well-being for this age group, and seeks to contribute to promoting mental health and supporting optimal care for individuals of very old age.

Methods

Study Design

We conducted a modified e-Delphi study to systematically combine expert opinions and achieve an informed group consensus on the symptomatology of depression in very old age. The Delphi technique is an established approach in which a panel of experts is asked to provide their opinions on a given issue over the course of several rounds [39]. In each round, the questions are informed by the findings of the previous one, allowing the study to evolve over time based on the collected data [39].

Given the prevalence of depression among older adults and the complexities of its diagnosis, the Delphi technique is particularly useful for gathering expert consensus on this issue. In our study, we used an e-Delphi approach to involve experts from different locations and facilitate efficient data collection through electronic communication methods. To adapt the traditional Delphi approach to the specific requirements of our study, we implemented several modifications (referred to as a *modified e-Delphi study*). To keep the focus on the items rather than on the degree of agreement, experts were informed at the start of each new round about which items had reached consensus without being given the precise consensus rates. In addition, for items on which consensus had already been reached, experts were unable to review their previous responses. This decision was made to avoid revisiting settled issues and manage the large number of items efficiently, ensuring the stability of agreed-upon elements. An additional aspect of our modified methodology was a final evaluation of open issues by the steering committee. This last step took place to discuss the items that had not achieved consensus over the e-Delphi rounds, ensuring a thorough examination of these elements.

The materials presented to the e-Delphi panel were based on a comprehensive analysis of the existing literature on depression in very old age (personal communication by Gomes da Rocha et al, 2023). This ensured that the expert panel's input was based on a thorough understanding of the relevant literature, thereby enhancing the validity of the final consensus, which may be seen as a lever to encourage scientific debate and foster new developments toward an effective diagnosis of depression in very old age [40].

This study adhered to the Conducting and Reporting Delphi Studies guidelines [41,42].

Steering Committee

The steering committee for this e-Delphi study based in Western Switzerland comprised 4 members: 2 physicians who are old age psychiatrists and senior researchers (AvG and PV), 1 clinical nurse specialist in geriatric care and a senior researcher skilled in the e-Delphi technique (HV), and 1 clinical nurse specialist in geriatric care who is also a junior researcher (CGdR). The committee was responsible for overseeing the study design, ensuring ethical compliance, and guiding the iterative process of questionnaire refinement and consensus building. Meetings were held monthly, with methodological decisions made based on discussion. The committee played a crucial role in maintaining the study's rigor, managing data collection and analysis, and addressing any lack of consensus among panelists.

Selection of Experts

In the Delphi approach, nonprobability sampling methods are used through the purposive selection of an expert panel. However, clear selection criteria were applied to limit researcher bias based on a five-step procedure [43]:

1. *Identifying the most appropriate categories of experts for the panel:* we identified experts from 3 main professional categories—physicians, nurses, and psychologists—each with a minimum of 5 years of professional experience and competencies in LLD. These professionals were selected for their specific work experience or contact with individuals of very old age in clinical, research, or teaching domains within the fields of old age psychiatry or geriatrics. In addition, our recruitment strategy aimed for a broad geographical representation by including experts from various continents (ie, Asia, Australia, Europe, and North America) to enrich the panel with diverse perspectives and, thereby, enhance the global view on depression screening among individuals of very old age.
2. *Identifying experts:* names were compiled from various sources, including previous research involvement, publications on the subject, professional email lists,

professional associations or societies, and boards of professional organizations. This method aimed to create a list of potential panelists recognized for their expertise and contributions to the field that was as comprehensive as possible.

3. *Contacting some readily reachable experts from each professional category to nominate other experts:* initial contact was made with experts who were easily accessible. These experts were then asked to nominate additional professionals whom they believed could contribute valuable insights to the study, thereby using a snowball sampling technique to ensure a wide and diverse pool of expertise.
4. *Creating sublists for each professional category and ranking experts based on specific criteria:* experts were organized into sublists by professional category and ranked according to their representation of professional role or specialty and their field of practice. This was a strategic approach intended to foster a diverse and balanced panel that reflects a broad spectrum of experiences and viewpoints, contributing to the robustness and credibility of the consensus process.
5. *Inviting the experts.*

The Delphi method offers no consensus on an ideal sample size as the focus is on qualitative depth and the use of expert judgment for a comprehensive exploration of specific topics rather than on statistical representativeness [44]. Usually, a panel comprises 7 to 15 experts to facilitate effective information processing and decision-making; however, some studies may involve up to 30 experts for broader perspectives [44]. For our study, we aimed to assemble a convenience sample of at least 20 experts from various countries.

Recruitment Procedure

On the basis of the predefined selection of potential panelists, we emailed invitation letters to a total of 84 experts. These letters outlined the study's objectives, described the e-Delphi context, detailed the questionnaire (eg, language, type of questions, and estimated completion time), and clarified what was expected from the participants, ensuring that they understood the significance of their contributions to the study. Interested parties could access the questionnaire through a direct URL to the survey. Anticipating a 25% to 30% response rate for the first-round questionnaire, we expected approximately 20 to 25 experts to participate. We aimed to retain these experts across all subsequent rounds, planning for a maximum of 4 rounds.

Data Collection Procedures

The SurveyMonkey tool (SurveyMonkey Inc) was used as the survey software as it facilitates secure and efficient data collection. In accordance with literature recommendations [43], each e-Delphi round remained open for a minimum of 4 weeks, during which up to 3 automatic reminder emails were dispatched to nonrespondents. At the start of each new round, experts were provided access to the overall results from the previous round, including a list of items for which consensus was either *achieved* (≥70% agreement) or *not achieved* (<70% agreement) [43]. The items that did not reach consensus were integrated into the current round for further deliberation, along with any new items that emerged from panel suggestions through open-ended questions.

Measures

The questionnaire was divided into 3 sections.

e-Delphi Section

This section aimed to achieve consensus on the topic under study and consisted of 104 items. These were divided into 24 dimensions (eg, *impaired concentration* and *lack of hope*) and 80 criteria (eg, *impaired ability to think or concentrate* and *hopelessness*), which experts rated according to their relevance using a Likert scale from 1 (*Not relevant*) to 4 (*Very relevant*). Experts also had the possibility to comment on their answers and suggest additional items that they considered relevant to the topic for subsequent rounds.

In addition to rating the relevance of each item, experts were asked to indicate the most suitable type of assessment for each criterion: self-assessment, hetero-assessment, or both. This was intended to identify preferred methods of assessing potential depressive signs and symptoms. The aim was not to achieve a stringent consensus of 70% as with the relevance ratings but to capture a broad spectrum of expert opinions on assessment approaches.

The initial selection of 104 items was based on a comprehensive review of existing instruments widely used in research to screen for and evaluate depressive signs and symptoms among near-centenarians and centenarians [13] supplemented by additional insights from recent analysis (personal communication by Gomes da Rocha et al, 2023). We started with all items available in these measures to be as comprehensive as possible. These instruments included self-assessment tools such as the Geriatric Depression Scale–30 [45], the Center for Epidemiologic Studies Depression Scale [46], the depression-related items from the Brief Symptom Inventory [47], and the Hospital Anxiety and Depression Scale [48], as well as the Montgomery-Åsberg Depression Rating Scale [49], which is administered by a professional rater. Furthermore, given the significant prevalence of neurocognitive disorders in our aging population and the need for reliable assessment of depressive symptoms in individuals with cognitive impairments, we also considered instruments specifically designed for such assessments: the Cornell Scale for Depression in Dementia [50,51] and the depression section of the Neuropsychiatric Inventory [52,53]. In addition, we included the *International Classification of Diseases, 11th Revision* and the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* diagnostic criteria to anchor our selections in the gold standards for depression diagnosis [6,54]. To complete the pool of items, the Depression Rating Scale from the Resident Assessment Instrument was also retained given its widespread use among adults of old and very old age in diverse settings, such as private homes (Resident Assessment Instrument–Home Care) or long-term care facilities (Resident Assessment Instrument–Long-Term Care Facilities) [17,55-57]. After compiling all items from the selected instruments, we checked for overlap and excluded identical items, which led to the final selection of 104 items. For instance, items such as "Do you feel

that your situation is hopeless?" from the Geriatric Depression Scale–30 and "Feeling hopeless about the future" from the Brief Symptom Inventory necessitated thematic grouping under a unified dimension (*Lack of Hope* in this example). This process enabled us to construct more general items designated as *dimensions* that reflected the essence of the criteria within each group, thereby streamlining the questionnaire without compromising the depth of depressive symptom screening. The list of all selected instruments and original items can be found in Multimedia Appendix 1.

In summary, this strategic selection of items aimed to cover a broad spectrum of relevant criteria for the screening of depressive signs and symptoms, ensuring the questionnaire's comprehensive scope and its applicability to adults of very old age with varying levels of cognitive functioning.

Sociodemographic and Professional Characteristics

This section included questions designed to capture the sociodemographic and professional characteristics of the panelists in round 1, including age, gender, nationality, profession, years of professional experience, and their country and field of practice. In subsequent rounds, only essential information (age, gender, nationality, and profession) was collected to ensure continuity and to accurately link responses from the same panelists across all rounds.

Contact With Individuals of Very Old Age

Although this could be considered part of professional characterization, this section specifically gathered information on the frequency of the panelists' experiences or interactions with individuals of very old age (aged ≥85 and ≥95 years). It was specific to the first round, aiming to inform the context of their responses. While this study's target population was defined as near-centenarians and centenarians (≥95 years), there is only a limited number of experts focusing on this exact population. However, most experts working with individuals aged ≥85 years have regular contact with near-centenarians and centenarians. Thus, these experts were also invited provided that they felt confident in drawing on their experiences and perceptions relevant to the population aged ≥95 years.

Pretest

Before launching round 1, the questionnaire was pretested with a small purposive sample of 5 professionals (nurses, psychologists, and a physician) to verify its clarity and comprehensibility, with only a minor adjustment made based on 1 expert's suggestion to include specific examples for certain items. Despite the positive outcomes of the pretest, experts were invited to rate the questionnaire at the end of round 1 for comprehensiveness, readability, and clarity.

Overall, 93% (26/28) of the experts participating in round 1 rated the questionnaire as *Comprehensive* or *Very comprehensive*, and 86% (24/28) found it *Easy* or *Very easy* to read and understand. Due to this positive feedback, no further structural adjustments were implemented before proceeding to the subsequent rounds.

Data Analysis

Data collected via SurveyMonkey were transferred to a Microsoft Excel spreadsheet (Microsoft Corp) and examined for any potential errors or missing data. They were then imported into the SPSS Statistics software (version 29.0; IBM Corp) for analysis [58]. Descriptive statistics were computed to characterize the panel from a sociodemographic and professional perspective using means and SDs as well as medians and interquartile ranges for quantitative variables and frequency distributions and percentages for qualitative variables. The level of agreement among experts required to achieve consensus on a given item was set at 70% for each round. This threshold is commonly used in Delphi studies [59-61] as it allows for the inclusion of minority viewpoints in the early rounds, ensuring that diverse perspectives are considered before achieving final consensus. This is particularly relevant in the context of depressive symptom screening given the challenges associated with detecting LLD, as previously mentioned. Hence, this consensus rate is considered acceptable for preserving methodological rigor, as supported by the literature [43], which indicates that such a threshold aligns with norms for ensuring reliability and validity in the context of an iterative, consensus-building approach. The consensus rate for each item was calculated using a 4-level Likert scale (1=*Not relevant*, 2=*Somewhat relevant*, 3=*Relevant*, and 4=*Very relevant*). In the first step, these responses were recoded into 2 levels: values of 1 and 2 were recoded as "0" (*Overall not relevant*), and values of 3 and 4 were recoded as "1" (*Overall relevant*). Next, we aggregated the recoded responses for each item to determine consensus, which was achieved when the percentage of experts identifying an item as *Overall relevant* or *Overall not relevant* met or exceeded the predefined threshold of 70%.

After the first e-Delphi round, items that did not reach the consensus level were carried forward to the subsequent round. At the beginning of each subsequent round, experts were given the opportunity to review the general outcomes from the previous round (ie, *consensus achieved* or *consensus not achieved* for each evaluated item). Items that achieved consensus were not included again in subsequent rounds. After conducting 3 rounds with the participating panel and observing only limited changes in the consensus rate after the third round, we decided to proceed with the last round, which involved a final evaluation of unresolved aspects by the steering committee. This last step aimed to gain a deeper understanding of the reasoning behind the outcomes and further refine the consensus process. Items that still did not achieve consensus after the 3 e-Delphi rounds were thoroughly examined during this phase.

Ethical Considerations

The Ethics Commission on Human Research of the Canton of Vaud was consulted and confirmed that this project fell outside the scope of the Swiss Human Research Act (Req-2023-00844; Multimedia Appendix 2) as the study did not involve the collection of health-related personal data or biological material.

All participants received detailed information about the study, including its aim and relevance, what to expect, a description of the questionnaire, instructions for participation, and contact information for the main researcher, along with the invitation

email. The questionnaires for all rounds (1, 2, and 3) included an introductory section emphasizing the voluntary nature of participation in the study, as well as specific information regarding confidentiality, data protection, risks, and benefits and a statement on ethical considerations. Participants were explicitly informed that, by answering the questionnaire, they were implicitly providing consent to take part in the study. They did not receive any compensation for their participation. Each participant received a survey link for accessing the questionnaire, and their IP addresses were not recorded to ensure confidentiality. The collected data were deidentified, with no personal identifiers retained. Responses stored on SurveyMonkey were deleted after transferring the data to the Microsoft Excel spreadsheet. All data are securely stored on a password-protected institutional server to comply with Swiss data protection regulations and will be definitively deleted after 5 years.

Results

Expert Participation and Sociodemographic and Professional Characteristics

A total of 84 experts were invited to this modified e-Delphi study. Of these 84 experts, 28 (33%) participated in the first round, forming the baseline panel. This group (n=28) was consequently invited to the second and third rounds, for which the participation rates were 75% (21/28) and 71% (20/28), respectively.

The panel exhibited sociodemographic and professional diversity, including different nationalities, professions, and fields of practice. The average age ranged slightly between 44.1 (SD 10.6) and 46.7 (SD 11.4) years over the rounds. A majority of women participated in the first and second rounds (17/28, 61% and 13/21, 62%, respectively), whereas men were more prevalent in the third round (11/20, 55%). Swiss and Portuguese nationalities were the most represented (12/20, 60% to 18/28, 64%), with the nursing profession being predominant (9/20, 45% to 12/21, 57%), followed by the medical profession (5/21, 24% to 8/20, 40%). The average professional experience was 19.4 (SD 8-9) years, with nearly 80% (22/28, 79%) of the experts being engaged in purely clinical activities or a combination of clinical and academic activities, among others (eg, service organization). More than 80% (23/28, 82%) reported having regular or very regular contact (daily, weekly, or monthly) with individuals aged ≥85 years, and 75% (21/28) reported having regular or very regular contact with individuals aged ≥95 years (Table 1).

Table 1. Sociodemographic and professional characteristics of the experts from the panel.

	Round 1 (n=28)	Round 2 (n=21)	Round 3 (n=20)
Age (y)			
Values, mean (SD)	44.1 (10.6)	45.2 (11.6)	46.7 (11.4)
Values, median (IQR)	39.5 (18.8)	39.0 (22.0)	44.0 (22.8)
Gender, n (%)			
Men	11 (39)	8 (38)	11 (55)
Women	17 (61)	13 (62)	9 (45)
Nationality, n (%)			
Portuguese	9 (32)	6 (29)	4 (20)
Swiss	9 (32)	7 (33)	8 (40)
French	3 (11)	2 (10)	3 (15)
Swiss and Portuguese	2 (7)	2 (10)	1 (5)
Australian	1 (4)	1 (5)	1 (5)
Belgian	1 (4)	1 (5)	1 (5)
German	1 (4)	1 (5)	0 (0)
Italian	1 (4)	0 (0)	1 (5)
Norwegian	1 (4)	1 (5)	1 (5)
Profession, n (%)			
Nurse	15 (54)	12 (57)	9 (45)
Physician	8 (29)	5 (24)	8 (40)
Psychologist	4 (14)	4 (19)	3 (15)
Other: gerontologist	1 (4)	0 (0)	0 (0)
Professional experience (y; n=27)			
Values, mean (SD)	19.4 (8.9)	__[a]	—
Values, median (IQR)	17.0 (11.0	—	—
Country of professional practice, n (%)			
Switzerland	15 (54)	—	—
Portugal	7 (25)	—	—
France	3 (11)	—	—
Australia	1 (4)	—	—
Norway	1 (4)	—	—
Portugal, United Kingdom, Saudi Arabia, and UAE[b]	1 (4)	—	—
Field of professional activity, n (%)			
Clinical activity			
Inpatient care only	6 (21)	—	—
Outpatient care only	2 (7)	—	—
Long-term care only	2 (7)	—	—
Academic activity			
Teaching only	1 (4)	—	—
Research only	1 (4)	—	—
Teaching and research	3 (11)	—	—
Clinical activity[c] and academic activity[d]	9 (32)	—	—

	Round 1 (n=28)	Round 2 (n=21)	Round 3 (n=20)
Clinical activity[c], academic activity[d], and other (eg, quality improvement or service organization)	3 (11)	—	—
Management activity	1 (4)	—	—
Contact with individuals aged ≥85 years, n (%)			
Daily	10 (36)	—	—
Weekly	9 (32)	—	—
Monthly	4 (14)	—	—
Yearly	4 (14)	—	—
Previously but not currently	1 (4)	—	—
Contact with individuals aged ≥95 years, n (%)			
Daily	4 (14)	—	—
Weekly	8 (29)	—	—
Monthly	9 (32)	—	—
Yearly (or less[e])	6 (21)	—	—
Previously but not currently	1 (4)	—	—

[a]Not applicable.

[b]UAE: United Arab Emirates.

[c]*Clinical activity* includes inpatient, outpatient, or long-term care settings.

[d]*Academic activity* includes teaching or research roles.

[e]One expert commented the following: "Not yearly, but once in a while."

Questionnaire Completion Times

The completion dates and the mean time taken for participants to complete the questionnaires in each round are summarized in Table 2. Globally, the e-Delphi rounds took place between August 2023 and February 2024, with an overall weighted mean completion time of 16 minutes and 40 seconds (SD 11 min and 27 s).

Table 2. Summary of overall consensus achievements across e-Delphi rounds regarding potential dimensions and criteria for depression screening.

	Questionnaire completion dates	Questionnaire completion time, mean (SD)	Dimensions analyzed	Consensus achieved (dimensions), n (%)	Consensus not achieved (dimensions), n (%)	New dimensions proposed	Criteria analyzed	Consensus achieved (criteria), n (%)	Consensus not achieved (criteria), n (%)	New criteria proposed
Round 1	August 15, 2023, to October 1, 2023	24 min, 51 s (10 min, 51 s)	24	13 (54)[a]	11 (46)[a]	4	80	49 (61)[b]	31 (39)[b]	8
Round 2	October 26, 2023, to December 12, 2023	15 min, 31 s (8 min, 18 s)	15	10 (67)[c]	5 (33)[c]	0	39	12 (31)[d]	27 (69)[d]	0
Round 3	January 31, 2024, to February 29, 2024	6 min, 26 s (3 min, 55 s)	5	1 (20)[e]	4 (80)[e]	0	27	9 (33)[f]	18 (67)[f]	0

[a]n=24.

[b]n=80.

[c]n=15.

[d]n=39.

[e]n=5.

[f]n=27.

Overall Consensus Achievements

A consensus threshold of ≥70% was established for both dimensions and criteria, as previously described.

Depression Screening: Potentially Relevant Dimensions

In round 1, experts assessed 24 dimensions, reaching consensus on 13 (54%). In addition, 4 new dimensions were suggested by the experts and incorporated in round 2: *underlying cognitive disorder* (dimension 25), *previous episodes of depression or diagnosed depression* (dimension 26), *personal care (eg, poor hygiene)* (dimension 27), and *language/speech (eg, poor speech)* (dimension 28). Thus, 15 dimensions (n=11, 73% without consensus and n=4, 27% new proposals) were reassessed in round 2, with consensus achieved on 10 (67%). In round 3, a total of 5 dimensions were revisited, with consensus reached on only 1 (20%; Table 2).

Table 3 summarizes the consensus rates for the 28 dimensions assessed across the 3 e-Delphi rounds. A total of 14% (4/28) of the dimensions failed to reach the ≥70% consensus threshold: memory problems (dimension 21), mood-congruent delusions (dimension 22), reduced adaptation capacity (dimension 24), and language/speech (eg, poor speech) (dimension 28)—the only newly suggested dimension without consensus. Of the dimensions on which consensus was achieved, all were considered relevant for screening for depression in individuals aged ≥95 years except for increased appetite (dimension 8), which was deemed not relevant. Notably, several relevant dimensions reached ≥90% consensus: lack of hope (dimension 20; 21/21, 100%), loss of interest (dimension 4; 27/28, 96%), lack of reactivity to pleasant events (dimension 5; 27/28, 96%), depressed mood (dimension 1; 26/28, 93%), and previous episodes of depression or diagnosed depression (dimension 26; 19/21, 90%).

Figure S1 in Multimedia Appendix 3 provides an overview of the consensus rates for the potentially relevant dimensions for depression screening.

Table 3. Consensus level on potential dimensions for depression screening[a].

	Round 1 (n=28)			Round 2 (n=21)			Round 3 (n=20)		
	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)
Depressed mood	Achieved	Relevant	26 (93)[b]	—[c]	—	—	—	—	—
Anxiety	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Irritability	Not achieved	N/A[d]	N/A	Achieved	Relevant	17 (81)[e]	—	—	—
Loss of interest	Achieved	Relevant	27 (96)[b]	—	—	—	—	—	—
Lack of reactivity to pleasant events	Achieved	Relevant	27 (96)[b]	—	—	—	—	—	—
Psychomotor changes	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Appetite loss	Achieved	Relevant	19 (70)[f]	—	—	—	—	—	—
Increased appetite	Not achieved	N/A	N/A	Achieved	Not relevant	16 (76)[e]	—	—	—
Reduced sleep or insomnia	Achieved	Relevant	19 (73)[g]	—	—	—	—	—	—
Hypersomnia	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Lack of energy	Achieved	Relevant	21 (78)[f]	—	—	—	—	—	—
Death wishes	Achieved	Relevant	22 (79)[b]	—	—	—	—	—	—
Suicidal thoughts	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Poor self-esteem	Achieved	Relevant	23 (82)[b]	—	—	—	—	—	—
Pessimism	Achieved	Relevant	20 (74)[f]	—	—	—	—	—	—
Low life satisfaction	Achieved	Relevant	20 (71)[b]	—	—	—	—	—	—
Impaired concentration	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Achieved	Relevant	14 (70)[h]
Impaired decision-making capacity	Not achieved	N/A	N/A	Achieved	Relevant	15 (71)[e]	—	—	—
Negative feelings	Achieved	Relevant	24 (86)[b]	—	—	—	—	—	—
Lack of hope	Not achieved	N/A	N/A	Achieved	Relevant	21 (100)[e]	—	—	—
Memory problems	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Mood-congruent delusions	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Repercussions of functional dependence	Not achieved	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Reduced adaptation capacity	Not achieved	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A
Underlying cognitive disorder (added in round 2)	N/A	N/A	N/A	Achieved	Relevant	15 (71)[e]	—	—	—
Previous episodes of depression or diagnosed depression (added in round 2)	N/A	N/A	N/A	Achieved	Relevant	19 (90)[e]	—	—	—

	Round 1 (n=28)			Round 2 (n=21)			Round 3 (n=20)		
	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)	Level of consensus at ≥70%	Result	Participants, n (%)
Personal care (eg, poor hygiene; added in round 2)	N/A	N/A	N/A	Achieved	Relevant	16 (76)[e]	—	—	—
Language/speech (eg, poor speech; added in round 2)	N/A	N/A	N/A	Not achieved	N/A	N/A	Not achieved	N/A	N/A

aThis list includes 4 dimensions suggested by the expert participants in round 1 and incorporated into the subsequent rounds: underlying cognitive disorder, previous episodes of depression or diagnosed depression, personal care (eg, poor hygiene), and language/speech (eg, poor speech).

[b]n=28.

[c]Data are not available for the corresponding entries.

[d]N/A: not applicable.

[e]n=21.

[f]n=27.

[g]n=26.

[h]n=20.

Depression Screening: Potentially Relevant Criteria

In round 1, a total of 80 criteria were assessed, with consensus reached on 49 (61%). In addition, 8 new criteria emerged from suggestions by the panel and were incorporated in round 2: *critical life events (eg, loss of a child or another loved one)* (criterion 81), *emotional indifference* (criterion 82), *anhedonia* (criterion 83), *afraid of being alone* (criterion 84), *resurfacing of "old wounds" (eg, long-standing conflicts with children)* (criterion 85), *feeling of accomplishment throughout life* (criterion 86), *closeness with family and significant others* (criterion 87), and *subjective health* (criterion 88). As a result, 39 criteria (n=31, 79% without consensus and n=8, 21% new proposals) were reassessed in round 2, with consensus achieved on 12 (31%). In round 3, a total of 27 criteria were revisited, with consensus reached on 9 (33%; Table 2).

The consensus rates for the 88 criteria assessed across the 3 e-Delphi rounds are presented in Multimedia Appendix 4. In total, 20% (18/88) of the criteria failed to reach the ≥70% consensus threshold: *easily annoyed* (criterion 20); *short-tempered* (criterion 21); *staying home instead of going out and doing new things* (criterion 26); *psychomotor agitation* (criterion 29); *significant unintentional weight gain (more than 5% in a month)* (criterion 33); *restless sleep* (criterion 35); *full of energy* (criterion 42); *feels as good as other people* (criterion 51); *pessimism* (criterion 52); *thinks most people are better off than him/her* (criterion 56); *finding life exciting, wonderful, and enjoyable* (criterion 58); *getting bored often* (criterion 62); *feeling fearful* (criterion 64); *feeling that people are unfriendly or dislike him/her* (criterion 66); *memory problems* (criterion 69); *the mind is as clear as it used to be* (criterion 70); *having a sense of direction and purpose in life* (criterion 79); and *afraid of being alone* (criterion 84)—the only newly suggested criterion without consensus. Of the criteria on which consensus was achieved, all were considered *relevant* for screening for depression in individuals aged ≥95 years except for *hard to get started on new projects* (criterion 28), *fear of dying* (criterion 75), and *sufficient financial resources* (criterion 77), which were considered *not relevant*. Notably, several *relevant* criteria reached a ≥90% consensus: *despondency, gloom, and despair* (criterion 5; 25/25, 100%); *depressed* (criterion 6; 27/27, 100%); *lack of reactivity to pleasant events or circumstances* (criterion 25; 28/28, 100%); *suicidal ideation* (criterion 45; 28/28, 100%); *suicide attempt(s)* (criterion 46; 28/28, 100%); *ruminations* (criterion 14; 27/28, 96%); *recurrent thoughts of death or suicide* (criterion 43; 27/28, 96%); *feelings of worthlessness* (criterion 47; 25/26, 96%); *critical life events* (criterion 81; 20/21, 95%); *anhedonia* (criterion 83; 20/21, 95%); *loss of interest in activities* (criterion 23; 26/28, 93%); *loss of pleasure in activities* (criterion 24; 26/28, 93%); and *sadness* (criterion 4; 24/26, 92%).

Figure S2 in Multimedia Appendix 3 provides an overview of the consensus rates for the potentially relevant criteria for depression screening.

Signs and Symptoms of Depression: Duration, Number, Frequency, and Severity

In round 1, experts (n=28) were asked about the importance of considering the duration of depressive signs or symptoms in addition to their presence. It is important to note that, for the purpose of this assessment, they were instructed not to consider *suicidal thoughts or ideation* as these require immediate attention and, thus, were deliberately excluded from the duration criteria. All participants (28/28, 100%) agreed on its importance. When specifically asked about the minimum duration for considering these signs or symptoms as clinically significant, 25 experts provided their insights—21 (84%) suggested a range from *a few days* (1 to 3) to *4 weeks*, with *2 weeks* being the most cited (n=9, 36%).

In addition, 7% (2/28) of the experts suggested differentiating the duration of signs or symptoms between recurrent episodes in chronic depression and new-onset cases. For this reason, a question addressing this specificity was introduced in round 2 (n=21), with 62% (13/21) supporting the distinction. For recurrent episodes, most experts favored a minimum duration

of *2 weeks* (9/13, 69%), whereas for a new onset, opinions varied, with "4 weeks" (5/13, 38%) being the most cited.

In round 2 (n=21), 62% (13/21) emphasized the importance of considering the number of signs or symptoms during screening. Of the 12 experts who specified a threshold, 7 (58%) favored a minimum of 3 symptoms.

Regarding sign or symptom frequency, all experts who provided an opinion in round 2 (20/20, 100%) considered it relevant. Of these 20 experts, 19 (95%) specified a particular frequency, of whom 9 (47%) suggested a frequency of *several times per day*, 3 (16%) suggested a frequency of *nearly daily*, and 6 (32%) suggested a frequency of *several times per week*.

For severity, 90% (18/20) considered it an important factor. Moreover, on a severity scale from 0 to 10 (with 10 being the most severe), the average threshold for clinical concern was 4.9 (SD 1.4), with a median of 5.0 (IQR 2.0).

Further details regarding the results described previously can be found in Multimedia Appendix 5.

Steering Committee's Decision on Dimensions and Criteria Lacking Consensus After 3 e-Delphi Rounds

The steering committee conducted 2 additional rounds of discussion to decide on the relevance of the 4 dimensions and 18 criteria that did not reach a consensus of ≥70% among the panelists during the 3 e-Delphi rounds.

In the first round of discussion, committee members reviewed the panelists' comments and agreement levels for each dimension and criterion, analyzing potential reasons for the lack of consensus and considering the clinical implications. Each member shared their perspectives drawing on their clinical expertise. In the second round, the committee revisited the dimensions and criteria after reflecting on the insights from the first round, aiming to resolve any remaining uncertainties and reach a collective decision. Consensus was prioritized, with majority voting used when necessary.

A summary of the committee's final determinations is provided in Multimedia Appendix 6, where the rationale behind the inclusion or exclusion of each dimension and criterion is detailed.

Table 4 presents an overview of the consensus levels reached in each e-Delphi round and the final decisions made by the steering committee.

Table 4. Steering committee's decision on the dimensions and criteria lacking consensus after 3 e-Delphi rounds.

	Round 1 (n=28), n (%)	Round 2 (n=21), n (%)	Round 3 (n=20), n (%)	Steering committee final decision
Dimensions				
Memory problems				Not relevant
Not relevant	15 (56)[a]	11 (52)[b]	9 (45)[c]	
Relevant	12 (44)[a]	10 (48)[b]	11 (55)[c]	
Mood-congruent delusions				Relevant
Not relevant	12 (44)[a]	9 (43)[b]	9 (45)[c]	
Relevant	15 (56)[a]	12 (57)[b]	11 (55)[c]	
Reduced adaptation capacity				Relevant
Not relevant	12 (43)[d]	7 (33)[b]	8 (40)[c]	
Relevant	16 (57)[d]	14 (67)[b]	12 (60)[c]	
Language or speech (eg, poor speech; added in round 2)				Not relevant
Not relevant	—[e]	8 (38)[b]	12 (60)[c]	
Relevant	—	13 (62)[b]	8 (40)[c]	
Criteria				
Easily annoyed				Not relevant
Not relevant	11 (39)[d]	12 (57)[b]	11 (55)[c]	
Relevant	17 (61)[d]	9 (43)[b]	9 (45)[c]	
Short-tempered				Relevant
Not relevant	11 (41)[a]	11 (58)[f]	10 (50)[c]	
Relevant	16 (59)[a]	8 (42)[f]	10 (50)[c]	
Staying home instead of going out and doing new things				Not relevant
Not relevant	11 (39)[d]	9 (43)[b]	12 (60)[c]	
Relevant	17 (61)[d]	12 (57)[b]	8 (40)[c]	
Psychomotor agitation				Not relevant
Not relevant	11 (41)[a]	8 (38)[b]	9 (45)[c]	
Relevant	16 (59)[a]	13 (62)[b]	11 (55)[c]	
Significant unintentional weight gain (more than 5% in a month)				Not relevant
Not relevant	18 (67)[a]	14 (67)[b]	11 (55)[c]	
Relevant	9 (33)[a]	7 (33)[b]	9 (45)[c]	
Restless sleep				Relevant
Not relevant	10 (38)[g]	8 (40)[c]	8 (40)[c]	
Relevant	16 (62)[g]	12 (60)[c]	12 (60)[c]	
Full of energy				Not relevant
Not relevant	18 (67)[a]	12 (57)[b]	12 (60)[c]	
Relevant	9 (33)[a]	9 (43)[b]	8 (40)[c]	
Feels as good as other people				Not relevant
Not relevant	14 (52)[a]	13 (65)[c]	8 (40)[c]	

	Round 1 (n=28), n (%)	Round 2 (n=21), n (%)	Round 3 (n=20), n (%)	Steering committee final decision
Relevant	13 (48)[a]	7 (35)[c]	12 (60)[c]	
Pessimism				Relevant
Not relevant	9 (32)[d]	9 (45)[c]	8 (40)[c]	
Relevant	19 (68)[d]	11 (55)[c]	12 (60)[c]	
Thinks most people are better off than him or her				Not relevant
Not relevant	10 (37)[a]	9 (43)[b]	7 (35)[c]	
Relevant	17 (63)[a]	12 (57)[b]	13 (65)[c]	
Finding life exciting, wonderful, and enjoyable				Not relevant
Not relevant	10 (37)[a]	9 (45)[c]	9 (45)[c]	
Relevant	17 (63)[a]	11 (55)[c]	11 (55)[c]	
Getting bored often				Not relevant
Not relevant	11 (41)[a]	12 (57)[b]	7 (37)[f]	
Relevant	16 (59)[a]	9 (43)[b]	12 (63)[f]	
Feeling fearful				Relevant
Not relevant	9 (33)[a]	11 (52)[b]	10 (50)[c]	
Relevant	18 (67)[a]	10 (48)[b]	10 (50)[c]	
Feeling that people are unfriendly or dislike him or her				Not relevant
Not relevant	15 (56)[a]	13 (62)[b]	7 (35)[c]	
Relevant	12 (44)[a]	8 (38)[b]	13 (65)[c]	
Memory problems				Not relevant
Not relevant	14 (50)[d]	12 (57)[b]	9 (45)[c]	
Relevant	14 (50)[d]	9 (43)[b]	11 (55)[c]	
The mind is as clear as it used to be				Not relevant
Not relevant	14 (52)[a]	11 (52)[b]	10 (50)[c]	
Relevant	13 (48)[a]	10 (48)[b]	10 (50)[c]	
Having a sense of direction and purpose in life				Relevant
Not relevant	11 (41)[a]	7 (33)[b]	9 (45)[c]	
Relevant	16 (59)[a]	14 (67)[b]	11 (55)[c]	
Afraid of being alone (added in round 2)				Relevant
Not relevant	—	9 (43)[b]	8 (40)[c]	
Relevant	—	12 (57)[b]	12 (60)[c]	

[a]n=27.
[b]n=21.
[c]n=20.
[d]n=28.
[e]Not applicable.
[f]n=19.
[g]n=26.

Types of Assessment for Potential Depression Criteria: Self-Assessment, Hetero-Assessment, or Both?

In addition to assessing the relevance of each criterion, the panel provided opinions on the preferred assessment method (self-assessment, hetero-assessment, or both). Rather than seeking strict consensus, the goal was to identify general trends. Using a 50% threshold as an example, 35% (28/80) of the criteria were deemed suitable for *self-assessment* only, 4% (3/80) of the criteria were deemed suitable for *hetero-assessment* only, and 40% (32/80) of the criteria were deemed suitable for both methods. Detailed information is available in Multimedia Appendix 7, whereas Figure S3 in Multimedia Appendix 3 provides an overview of the consensus rates for these assessment methods.

Discussion

Principal Findings

This modified e-Delphi study aimed to identify and achieve consensus on the most representative signs and symptoms of depression in very old age, focusing on near-centenarians and centenarians.

Our results revealed that most of the items proposed—86% (24/28) of the dimensions and 80% (70/88) of the criteria—which had been assembled based on a literature review, were considered *relevant* for depression screening in very old age. This highlights the importance of a multidimensional approach and acknowledges the complexity of accurately diagnosing depression in this population.

Most Consensual Potentially Relevant Depressive Signs and Symptoms

Focusing on items with the highest consensus rates (≥90%), some overlap between dimensions and criteria was observed, which was expected as dimensions grouped related criteria.

For example, the *depressed mood* dimension was widely considered *relevant*, aligning with criteria such as *depressed*; *sadness*; and *despondency, gloom, and despair*, which reflect the emotional core of depressive disorders [6,54]. Similarly, *loss of interest in activities*, *loss of pleasure in activities*, *lack of reactivity to pleasant events or circumstances*, and *anhedonia* also achieved broad consensus as critical indicators of depression [6,54,62-64]. Collectively, these signs and symptoms highlight the significant impact of depression on a person's capacity for joy and engagement with life, underscoring their relevance in depression screening.

Criteria such as *ruminations* and *feelings of worthlessness* equally achieved a strong consensus due to their role in the negative cognitive and emotional patterns of depression. Previous research has shown that *ruminations* exacerbate psychopathology, including depression, by intensifying and extending negative mood states, thus interfering with problem-solving and perpetuating physiological stress responses [65]. Similarly, *feelings of worthlessness* are strongly linked to self-blame, lack of hope, and an increased suicide risk even after depressive episodes remit [66,67]. Building on this, the *lack ofhope* dimension achieved unanimous consensus (21/21, 100%) as a relevant element, reflecting its importance as a significant predictor of psychological well-being [68]. Within this broader dimension, the *hopelessness* and *hopeful about the future* criteria also showed high consensus (24/27, 89% and 16/21, 76%, respectively), emphasizing their important role in depression screening in individuals of very old age whether as a negative indicator or a positive resource.

Furthermore, *previous episodes of depression or diagnosed depression* were also identified as relevant, highlighting the importance of considering an individual's psychiatric history as past experiences with depression increase the risk of recurrence [69]. In addition, *critical life events* perceived as negative were recognized for their significant association with higher levels of depressive symptoms in a recent meta-analysis [70].

Finally, the strong agreement on severe criteria such as *recurrent thoughts of death or suicide*, *suicidal ideation*, and past *suicide attempt(s)* underscores the critical need to address these factors due to their immediate risk to individuals' safety [26].

Most Consensual Potentially Relevant Depressive Signs and Symptoms in the Context of Old Age

Having established the most consensual depressive features, it is essential to explore how these characteristics manifest specifically in older adults as unique age-related factors may shape their expression and impact.

As individuals age, they may face different challenges, some of which are particularly relevant to LLD. These include agism; loss of income; the death of loved ones; and the risk of isolation, functional limitations, multimorbidity, cognitive decline, and institutionalization [71]. Such multidimensional challenges can impact well-being and mental health, contributing to reduced aging satisfaction and increased loneliness, including near-centenarians and centenarians [71,72]. While the cumulative effect of these difficulties can increase the risk of LLD, it is important to note that experiencing them does not necessarily lead to the development of a depressive state [72].

The dimensions and criteria that achieved high consensus, as previously discussed, clearly represent central elements in depression screening across all age groups, aligning with current diagnostic standards [6,54]. Our findings confirm their importance for accurate diagnosis in individuals of very old age, underscoring their transversality across generations and relevance even in the final stages of life. While depression in very old age may have unique aspects, some of the core signs and symptoms of the disorder remain consistent and should not be overlooked, including in near-centenarians and centenarians. However, these signs and symptoms may carry additional significance due to the unique psychosocial and physical challenges faced by these individuals [72].

For instance, *loss of interest in activities*, *loss of pleasure in activities*, *lack of reactivity to pleasant events or circumstances*, and *anhedonia* may highlight a deeper disengagement from life that can be difficult to detect in older adults. These symptoms are often masked by typical aging challenges such as physical decline, social isolation, or even cognitive issues, complexifying their identification. In addition, *feelings of worthlessness* and

the negative impact of *critical life events* may be particularly relevant in late-life contexts, where significant life changes and losses are common [67,73]. As highlighted in a meta-synthesis [74], older adults often experience declines in physical and cognitive abilities, leading to feelings of disconnection and a perception of being unneeded by their communities. This sense of being unneeded can foster feelings of isolation and insignificance, thereby triggering or exacerbating depressive symptoms. Moreover, societal changes and loss of social roles further deepen these feelings of worthlessness, underscoring the importance of addressing these issues in older populations [74].

While the high consensus on many classic criteria suggests their continued relevance across the life cycle, our findings also point to unique aspects of depression in very old age that may not be fully captured by current approaches. For example, the unanimous agreement on the relevance of *lack ofhope* highlights its particular significance in this age group. Unlike other factors typically associated with normal aging processes, low levels of hope have been identified as a significant risk factor in older adults, reflecting a deeper sense of despair with potentially serious consequences. A systematic review [75] underscored that lack of hope is closely linked to negative outcomes in older populations, including an increased risk of depression and poorer overall health outcomes.

Moreover, the strong agreement on severe symptoms such as *recurrent thoughts of death or suicide*, *suicidal ideation*, and past *suicide attempt(s)* underscores the importance of also addressing these factors in very old age, where such topics often remain taboo [76]. Reluctance to discuss or acknowledge these symptoms may lead to underdiagnosis and inadequate treatment, emphasizing the need for sensitive and proactive approaches in this age group. Recent qualitative research [76] has revealed that suicidal thoughts in later life are often subtly expressed and normalized in everyday conversations, making them difficult to detect. However, the extent to which this applies to centenarians is unclear. Previous research suggests that only a minority of centenarians express a wish to die [77,78], making explicit suicidal thoughts likely rare and not always linked to depressive states. Nonetheless, feelings of life being *completed and no longer worth living* reflect a deeper disconnection from life, characterized by aching loneliness, feelings of not mattering, and a physical and mental tiredness [79]. Creating safe spaces for informal—yet meaningful—discussions is essential, alongside strategies for timely recognition and management of these symptoms [76].

Potentially Not Relevant Depressive Signs and Symptoms in the Context of Old Age

The *increased appetite* dimension may have been considered *not relevant* as it can result from situational factors such as attending a day center or moving to a nursing home, where tailored meals and social dining often enhance the willingness to eat, contributing to an increased appetite. Similarly, the *significant unintentional weight gain* criterion was likely deemed *not relevant* for related reasons as previous research has found no association between weight gain and incident depressive symptoms in older adults [80].

Certain cognitive dimensions such as *memory problems* and *language/speech (eg, poor speech)* were also classified as *not relevant*, likely due to their overlap with nondepressive phenomena commonly observed in older adults. This highlights the need for tailored screening protocols that can effectively differentiate between depression and cognitive impairment, conditions that often coexist in individuals of very old age and require distinct therapeutic strategies [81,82]. The criterion *the mind is as clear as it used to be* was likely deemed *not relevant* for the same reasons.

Criteria such as *hard to get started on new projects*; *staying home instead of going out and doing new things*; *finding life exciting, wonderful, and enjoyable*; *easily annoyed*; and *getting bored often* may be more closely related to motivational factors, social involvement, social support, or underlying health conditions rather than directly indicating a depressive state [83,84]. Similarly, criteria such as *thinks most people are better off than him/her*, *feeling that people are unfriendly or dislike him/her*, and *feels as good as other people* may reflect personality traits, which are relatively stable over the life span and may not necessarily reflect a change in mental health status [85].

Symptoms such as *psychomotor agitation* illustrate the challenge in differentiating between depression-related symptoms and those arising from other geriatric conditions such as dementia or delirium [86], suggesting that certain depressive features may be less applicable or not sufficient for individuals of very old age [87]. In addition, the *full of energy* criterion, typically seen as a positive indicator, was not deemed relevant, possibly because it is rare for a near-centenarian or centenarian without depression to report feeling full of energy. This likely reflects their overall physical condition rather than a depressive state, and the absence of this criterion may often be confounded with physical exhaustion linked to frailty [88].

The *fear of dying* criterion was also classified as *not relevant*. Although it might be assumed that individuals in the final stage of life would fear death, previous research suggests that death anxiety is generally low among older adults, who are more often concerned with the dying process rather than death itself [89]. Similarly, *sufficient financial resources* was deemed *not relevant*. While financial difficulties can impact mental health, their link to psychological distress in older adults appears to be more influenced by underlying factors such as mastery and self-esteem rather than acting as a direct indicator of depression [90].

The Specificity of Near-Centenarians and Centenarians

Research on near-centenarians and centenarians has predominantly focused on physical, cognitive, and social health, with less attention paid to psychiatric aspects [13]. Studies on depression in this population often rely on overall assessment scores from standardized instruments rather than detailed analysis of individual items [72,87,91-94]. This gap, alongside the variety of existing tools for screening for depression in very old age, also motivated this study.

Notably, one study conducted a detailed item-level analysis using a 14-item version of the Geriatric Depression Scale in

Portuguese centenarians [88] and found that 51.4% reported *feelings of worthlessness*, supporting the relevance of this specific criterion in our e-Delphi study. In addition, only 12.3% reported having more *memory problems*, 43.3% expressed a preference for *staying home instead of going out and doing new things*, 39.8% reported *getting bored often*, and 33.3% agreed with the statement *thinks most people are better off than they are* [88]. Globally, these findings align with our results, where these criteria were considered not relevant in this age group.

While these findings suggest reduced presence of symptoms established for younger populations, it remains unclear whether depression in near-centenarians and centenarians significantly differs from that in other older individuals, such as octogenarians; the key features of depression may not vary meaningfully. Therefore, it is worth questioning whether the development of a tool specifically designed for very old age, including the near-centenarian and centenarian population, is truly necessary. Nevertheless, psychological resources such as resilience appear to be particularly strong in near-centenarians and centenarians [77,95]. These protective characteristics may lead traditional depression screening tools to overlook subtle signs of depression that are masked by these strengths. A tool specifically designed for this age group could better capture the subtle differences and specific needs, enhancing the accuracy and effectiveness of depression screening.

Duration, Number, Frequency, and Severity of Depressive Signs and Symptoms and Type of Assessment

Our findings emphasize the importance of considering the duration, number, frequency, and severity of depressive signs and symptoms in near-centenarians and centenarians. Notably, experts differentiated between chronic and new-onset depression, advocating for a nuanced approach that may significantly impact clinical assessments and interventions. This aligns with existing literature documenting distinct clinical profiles for persistent and nonchronic depressive disorders in adults aged 18 to 79 years [96,97], with persistent cases often involving earlier onset, more severe symptoms, and greater treatment resistance [96]. While these findings are based on younger individuals, they reinforce the necessity of a personalized life span approach that includes rigorous screening and comprehensive clinical strategies for the oldest age groups.

Furthermore, experts in our study highlighted the potential value of combining self- and hetero-assessment techniques to screen for depressive signs and symptoms in very old age. However, this approach may require further consideration. Cognitive limitations at this age may affect the reliability of self-reports, complicating the integration with caregiver observations. Developing strategies to reconcile these differing perspectives—such as guidelines for interpreting discrepancies—could ensure that assessments are accurate and sensitive to the unique challenges faced by individuals of very old age. Flexible, combined approaches remain relevant as research suggests that they enhance diagnostic accuracy and patient engagement [98].

Further Considerations

Given the findings of our study, a more comprehensive approach to depression screening in near-centenarians and centenarians may be advantageous. The high consensus among experts regarding certain depressive features highlights their relevance and significance for this particular age group. Importantly, the diagnosis of LLD in adults of very old age can be particularly complex due to overlapping symptoms with other mental or somatic conditions, as previously outlined. This overlap necessitates a differential diagnostic approach that integrates a broader perspective. This emphasizes the need for screening tools that are both comprehensive and sensitive to the unique experiences and needs of individuals of very old age. Implementing such tailored screening methods could contribute to improved mental health outcomes within this population.

Our work extends existing literature by providing a comprehensive overview of a wider range of depressive features, including considerations of their duration, number, frequency, and severity—factors that are less considered in standard depression screenings. This highlights the complexity of diagnosing depression in adults of very old age and confirms the necessity of incorporating these specific aspects into the screening process.

While the e-Delphi method was valuable for identifying key elements in detecting depressive signs and symptoms in those of very old age, a separate validation study is necessary to confirm whether the extended screening process would be more effective compared to current tools. Nonetheless, this study hopefully represents a meaningful step forward in improving the detection of depressive signs and symptoms in this population.

Limitations

We acknowledge several limitations. First, despite the international and professional diversity of the expert panel, it may not have captured the full spectrum of clinical practices or cultural perceptions related to aging and depression. The relative overrepresentation of nurses among the panelists in rounds 1 and 2 may have introduced a bias toward nursing-specific priorities, such as a stronger focus on the practical applicability of screening criteria in daily care. However, this also reflects their pivotal role in direct care for individuals of very old age, making their insights particularly relevant for developing pragmatic screening criteria. Given this, a *groupthink* effect in which the dominant profession influenced the consensus cannot be excluded. In addition, our panel lacked experts from Asian countries despite notable recruitment efforts even though these regions have some of the highest proportions of centenarians worldwide. This absence may have introduced a bias toward Western perspectives. At the same time, the concept of depression in very old age may be less clearly defined in some Asian cultures due to cultural background. Second, the selection of a 70% consensus threshold may have influenced the interpretation of the results as it reflects a compromise between inclusivity and stringency, subtly shaping the scope of the consensus achieved. Third, the structured nature of the e-Delphi process may have limited the exploration of nuanced symptoms critical to understanding depression in individuals of very old

age, a population whose cognitive decline may complicate symptom presentation. Fourth, some overlap between dimensions and criteria was observed, which may be perceived as a limitation to some extent. However, the steering committee was aware of this overlap and intentionally decided to retain it to confirm certain trends in consensus on specific items.

Despite these challenges, this study's insights have the potential to stimulate discussions on improving mental health strategies for older populations, especially those of very old age. The diverse and interdisciplinary composition of the panel further contributes to the generalizability of the findings, offering valuable insights to health care professionals and researchers worldwide.

Future research should aim to develop and validate a screening tool specifically designed for depressive signs and symptoms in very old age, including near-centenarians and centenarians. A validation study comparing the extended method with established tools or a *gold standard* diagnostic reference is crucial to determining its efficacy. In addition, it is important to assess the tool's effectiveness across different cultural contexts to ensure its broad applicability. The incorporation of Patient-Reported Experience Measures in future developments would also be beneficial in capturing the lived experiences and perceptions of this age group.

Conclusions

This expert consensus could significantly contribute to enhancing early detection and intervention for depression in individuals of very old age, notably by informing the creation of a new screening tool. The findings have the potential to improve diagnostic accuracy and, consequently, allow for the personalization of treatment plans tailored to the unique needs of individuals of very old age.

In conclusion, this study used expert consensus to establish a foundational step toward enhancing the detection and treatment of depression among individuals of very old age, including near-centenarians and centenarians. This study emphasizes the need for cautious, multidimensional, and multidisciplinary approaches to depression screening in this age group, advocating for the continuous refinement of screening tools to address the complex nature of very old age effectively.

Acknowledgments

This work was supported by the Swiss National Science Foundation through the Sinergia grant CRSII5_186239/1 (SWISS100). The authors gratefully acknowledge the contributions of all the experts who participated in this study. They recognize those who agreed to be publicly acknowledged: Alcina Matos Queirós, Department of Health and Social Welfare, Canton of Vaud, Switzerland; Carina Santos, Riviera-Chablais Hospital, Switzerland; Christophe Nakamura, Lausanne University Hospital, Switzerland; Eivind Aakhus, the Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Norway; Filipa Pereira, School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais-Wallis, Switzerland; Jean Roche, Lille University Hospital, France; Kim Uittenhove, Swiss Federal Institute of Technology in Lausanne, Switzerland; Nancy A Pachana, the University of Queensland, Australia; Oscar Ribeiro, University of Aveiro, Portugal; Raquel Neves, Almoosa Rehabilitation Hospital, Kingdom of Saudi Arabia; Sofia Fernandes, School of Health Sciences, University of Applied Sciences and Arts Western Switzerland Valais-Wallis, Switzerland; Sylvain Nguyen, Riviera-Chablais Hospital, Switzerland; Sylvie Bonin-Guillaume, Aix-Marseille University, University Hospital of Marseille, France; and Tiago Ferreira, Lausanne University Hospital, Switzerland. The authors also extend their sincere appreciation to the experts who preferred to remain anonymous. Their equally valuable contributions were central in shaping this study's outcomes.

Authors' Contributions

CGdR, AvG, and HV contributed to the conceptualization of the study. AvG, PV, and HV provided methodological input. AvG, PV, DSJ, OR, and HV contributed scientific input. CGdR was responsible for data collection, data analysis, and writing the original draft. CGdR, AvG, PV, DSJ, OR, and HV reviewed and edited the manuscript. AvG and HV supervised the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Selected instruments and their respective items.
[DOCX File , 27 KB - aging_v8i1e64352_app1.docx]

Multimedia Appendix 2
Decision of the ethics committee.
[DOCX File , 180 KB - aging_v8i1e64352_app2.docx]

Multimedia Appendix 3
Graphical overview of the consensus rates (dimensions, criteria, and type of assessment).

[DOCX File , 774 KB - aging_v8i1e64352_app3.docx]

Multimedia Appendix 4
Consensus level on potential criteria for depression screening.
[DOCX File , 42 KB - aging_v8i1e64352_app4.docx]

Multimedia Appendix 5
Aspects of signs and symptoms of depression (duration, number, frequency, and severity).
[DOCX File , 20 KB - aging_v8i1e64352_app5.docx]

Multimedia Appendix 6
The steering committee's final determinations.
[DOCX File , 17 KB - aging_v8i1e64352_app6.docx]

Multimedia Appendix 7
Findings on the preferred assessment methods for depression screening criteria.
[DOCX File , 25 KB - aging_v8i1e64352_app7.docx]

References

1. Schaeffer K. U.S. centenarian population is projected to quadruple over the next 30 years. Pew Research Center. URL: https://tinyurl.com/4z8tuu3d [accessed 2024-04-08]
2. World population ageing. United Nations (UN). URL: https://tinyurl.com/33chm345 [accessed 2024-04-29]
3. The world population prospects: 2015 revision. United Nations (UN). URL: https://www.un.org/en/development/desa/publications/world-population-prospects-2015-revision.html [accessed 2024-04-29]
4. Teixeira L, Araújo L, Jopp D, Ribeiro O. Centenarians in Europe. Maturitas 2017 Oct;104:90-95. [doi: 10.1016/j.maturitas.2017.08.005] [Medline: 28923181]
5. Living to 100 years and beyond. Federal Statistical Office. URL: https://tinyurl.com/bdzcvhzd [accessed 2024-04-08]
6. ICD-11 for mortality and morbidity statistics. World Health Organization. URL: https://icd.who.int/browse/2024-01/mms/en [accessed 2024-04-29]
7. Rapport mondial sur le vieillissement et la santé. World Health Organization. URL: https://tinyurl.com/mrwmk5bs [accessed 2024-04-29]
8. Jopp DS, Boerner K, Rott C. Health and disease at age 100: findings from the second Heidelberg Centenarian studies. Dtsch Arztebl Int 2016 Mar 25;113(12):203-210 [FREE Full text] [doi: 10.3238/arztebl.2016.0203] [Medline: 27118718]
9. Devita M, de Salvo R, Ravelli A, de Rui M, Coin A, Sergi G, et al. Recognizing depression in the elderly: practical guidance and challenges for clinical management. Neuropsychiatr Dis Treat 2022;18:2867-2880 [FREE Full text] [doi: 10.2147/NDT.S347356] [Medline: 36514493]
10. Lleshi V, Bizzozzero T. La dépression du sujet âgé. Rev Med Suisse 2009;5:1785-1789. [doi: 10.53738/revmed.2009.5.216.1785]
11. Cappeliez P. La dépression. In: Voyer P, editor. Soins infirmiers aux aînés en perte d'autonomie 3e. Montréal, Québec: Éditions du Renouveau Pédagogique Inc; 2021:185-198.
12. Gallarda T, Lôo H. Dépression et personnes âgées. Encephale 2009 Jun;35(3):269-280. [doi: 10.1016/j.encep.2008.10.013] [Medline: 19540414]
13. Cheng A, Leung Y, Harrison F, Brodaty H. The prevalence and predictors of anxiety and depression in near-centenarians and centenarians: a systematic review. Int Psychogeriatr 2019 Nov 29;31(11):1539-1558 [FREE Full text] [doi: 10.1017/S1041610219000802] [Medline: 31354113]
14. Eshun S, Gurung RA. Culture and Mental Health: Sociocultural Influences, Theory, and Practice. Halifax, NS: Wiley-Blackwell; 2009.
15. Bergdahl E, Gustavsson JM, Kallin K, von Heideken Wågert P, Lundman B, Bucht G, et al. Depression among the oldest old: the Umeå 85+ study. Int Psychogeriatr 2005 Dec;17(4):557-575 [FREE Full text] [doi: 10.1017/S1041610205002267] [Medline: 16185377]
16. Clément JP. La dépression à l'âge avancé. In: Giannakopoulos P, Gaillard M, editors. Abrégé de Psychiatrie de l'Âge Avancé. Chêne-Bourg, Switzerland: Médecine & Hygiène; 2010:171-183.
17. Hirdes JP, Ikegami N, Jónsson PV, Topinková E, Maxwell CJ, Yamauchi K. Cross-national comparisons of antidepressant use among institutionalized older persons based on the minimum data set (MDS). Can J Aging 2010 Nov 29;19(S2):18-37. [doi: 10.1017/S0714980800013878]
18. Hottin P, Trudel JF. Dépression et toubles affectifs. In: Arcand M, Hébert R, editors. Précis Pratique de Gériatrie. Montreal, QC: Edisem inc; 2007:305-322.

19. Mental health of older adults. World Health Organization (WHO). 2017. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2024-04-29]
20. Wei J, Hou R, Zhang X, Xu H, Xie L, Chandrasekar EK, et al. The association of late-life depression with all-cause and cardiovascular mortality among community-dwelling older adults: systematic review and meta-analysis. Br J Psychiatry 2019 Aug;215(2):449-455. [doi: 10.1192/bjp.2019.74] [Medline: 30968781]
21. Cherbuin N, Kim S, Anstey KJ. Dementia risk estimates associated with measures of depression: a systematic review and meta-analysis. BMJ Open 2015 Dec 21;5(12):e008853 [FREE Full text] [doi: 10.1136/bmjopen-2015-008853] [Medline: 26692556]
22. Diniz BS, Butters MA, Albert SM, Dew MA, Reynolds 3rd CF. Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. Br J Psychiatry 2013 May;202(5):329-335 [FREE Full text] [doi: 10.1192/bjp.bp.112.118307] [Medline: 23637108]
23. John A, Patel U, Rusted J, Richards M, Gaysina D. Affective problems and decline in cognitive state in older adults: a systematic review and meta-analysis. Psychol Med 2019 Feb;49(3):353-365 [FREE Full text] [doi: 10.1017/S0033291718001137] [Medline: 29792244]
24. Santabarbara J, Sevil-Perez A, Olaya B, Gracia-Garcia P, Lopez-Anton R. [Clinically relevant late-life depression as risk factor of dementia: a systematic review and meta-analysis of prospective cohort studies]. Rev Neurol 2019 Jun 16;68(12):493-502 [FREE Full text] [doi: 10.33588/rn.6812.2018398] [Medline: 31173329]
25. Thomas A. Depression in older people. In: Dening T, Thomas A, editors. Oxford Textbook of Old Age Psychiatry. 2nd edition. New York, NY: Oxford University Press; 2013:545-569.
26. Gramaglia C, Calati R, Zeppegno P. Rational suicide in late life: a systematic review of the literature. Medicina (Kaunas) 2019 Sep 29;55(10):52 [FREE Full text] [doi: 10.3390/medicina55100656] [Medline: 31569542]
27. Bian Z, Xu R, Shang B, Lv F, Sun W, Li Q, et al. Associations between anxiety, depression, and personal mastery in community-dwelling older adults: a network-based analysis. BMC Psychiatry 2024 Mar 07;24(1):192 [FREE Full text] [doi: 10.1186/s12888-024-05644-z] [Medline: 38454373]
28. Balsamo M, Cataldi F, Carlucci L, Padulo C, Fairfield B. Assessment of late-life depression via self-report measures: a review. Clin Interv Aging 2018;13:2021-2044 [FREE Full text] [doi: 10.2147/CIA.S178943] [Medline: 30410319]
29. Luppa M, Sikorski C, Motzek T, Konnopka A, König HH, Riedel-Heller SG. Health service utilization and costs of depressive symptoms in late life - a systematic review. Curr Pharm Des 2012;18(36):5936-5957. [doi: 10.2174/138161212803523572] [Medline: 22681171]
30. British Geriatrics Society. Depression among older people living in care homes report. Royal College of Psychiatrists and British Geriatrics Society. URL: https://tinyurl.com/5e4tx7yj [accessed 2024-04-29]
31. Sudhir Kumar CT, George S, Augustine A. Treatment of late-life depression. J Geriatr Care Res 2020;7(1):9-16 [FREE Full text]
32. Fußer F, Karakaya T, Pantel J. Late-life depression. In: Carvalho AF, McIntyre RS, editors. Mental Disorders in Primary Care: A Guide to their Evaluation and Management. New York, NY: Oxford University Press; 2017:251-264.
33. Charoenngam N, Rittiphairoj T, Ponvilawan B, Prasongdee K. Thyroid dysfunction and risk of Parkinson's disease: a systematic review and meta-analysis. Front Endocrinol (Lausanne) 2022;13:863281 [FREE Full text] [doi: 10.3389/fendo.2022.863281] [Medline: 35600588]
34. Dagnino AP, Campos MM. Chronic pain in the elderly: mechanisms and perspectives. Front Hum Neurosci 2022;16:736688 [FREE Full text] [doi: 10.3389/fnhum.2022.736688] [Medline: 35308613]
35. Lee PH, Yeh TT, Yen HY, Hsu WL, Chiu VJ, Lee SC. Impacts of stroke and cognitive impairment on activities of daily living in the Taiwan longitudinal study on aging. Sci Rep 2021 Jun 09;11(1):12199 [FREE Full text] [doi: 10.1038/s41598-021-91838-4] [Medline: 34108582]
36. Liu M, Hou T, Nkimbeng M, Li Y, Taylor JL, Sun X, et al. Associations between symptoms of pain, insomnia and depression, and frailty in older adults: a cross-sectional analysis of a cohort study. Int J Nurs Stud 2021 May;117:103873 [FREE Full text] [doi: 10.1016/j.ijnurstu.2021.103873] [Medline: 33621722]
37. Pazan F, Wehling M. Polypharmacy in older adults: a narrative review of definitions, epidemiology and consequences. Eur Geriatr Med 2021 Jun 10;12(3):443-452 [FREE Full text] [doi: 10.1007/s41999-021-00479-3] [Medline: 33694123]
38. Wang EY, Meyer C, Graham GD, Whooley MA. Evaluating screening tests for depression in post-stroke older adults. J Geriatr Psychiatry Neurol 2018 May;31(3):129-135. [doi: 10.1177/0891988718778791] [Medline: 29793370]
39. Barrett D, Heale R. What are Delphi studies? Evid Based Nurs 2020 Jul;23(3):68-69. [doi: 10.1136/ebnurs-2020-103303] [Medline: 32430290]
40. Cole ZD, Donohoe HM, Stellefson ML. Internet-based Delphi research: case based discussion. Environ Manage 2013 Mar;51(3):511-523 [FREE Full text] [doi: 10.1007/s00267-012-0005-5] [Medline: 23288149]
41. Spranger J, Homberg A, Sonnberger M, Niederberger M. Reporting guidelines for Delphi techniques in health sciences: a methodological review. Z Evid Fortbild Qual Gesundhwes 2022 Aug;172:1-11 [FREE Full text] [doi: 10.1016/j.zefq.2022.04.025] [Medline: 35718726]

42. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on conducting and REporting DElphi studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 2017 Sep 13;31(8):684-706. [doi: 10.1177/0269216317690685] [Medline: 28190381]
43. Gill FJ, Leslie GD, Grech C, Latour JM. Using a web-based survey tool to undertake a Delphi study: application for nurse education research. Nurse Educ Today 2013 Nov;33(11):1322-1328. [doi: 10.1016/j.nedt.2013.02.016] [Medline: 23510701]
44. Naisola-Ruiter V. The Delphi technique: a tutorial. Res Hosp Manag 2022 Jun 06;12(1):91-97. [doi: 10.1080/22243534.2022.2080942]
45. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
46. Radloff LS. The CES-D scale. Applied Psychol Manag 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
47. Derogatis L. Brief Symptom Inventory (BSI): Administration, Scoring, and Procedures Manual. 4th edition. Minneapolis, MN: National Computer Systems; 1993.
48. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
49. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
50. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell scale for depression in Dementia. Biol Psychiatry 1988 Feb 01;23(3):271-284 [FREE Full text] [doi: 10.1016/0006-3223(88)90038-8] [Medline: 3337862]
51. Resnick B, van Haitsma K, Kolanowski A, Galik E, Boltz M, Ellis J, et al. Reliability and validity of the Cornell Scale for depression in dementia and invariance between black versus white residents in nursing homes. J Am Med Dir Assoc 2022 Jul;23(7):1236-41.e3 [FREE Full text] [doi: 10.1016/j.jamda.2021.11.016] [Medline: 34896057]
52. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994 Dec;44(12):2308-2314. [doi: 10.1212/wnl.44.12.2308] [Medline: 7991117]
53. Kaufer DI, Cummings JL, Christine D, Bray T, Castellon S, Masterman D, et al. Assessing the impact of neuropsychiatric symptoms in Alzheimer's disease: the Neuropsychiatric Inventory Caregiver Distress Scale. J Am Geriatr Soc 1998 Feb;46(2):210-215. [doi: 10.1111/j.1532-5415.1998.tb02542.x] [Medline: 9475452]
54. American Psychiatric Association. Diagnostic and Statistical Manual Of Mental Disorders: DSM-5™, 5th edition. New York, NY: American Psychiatric Publishing; 2013.
55. Burrows AB, Morris JN, Simon SE, Hirdes JP, Phillips C. Development of a minimum data set-based depression rating scale for use in nursing homes. Age Ageing 2000 Mar;29(2):165-172. [doi: 10.1093/ageing/29.2.165] [Medline: 10791452]
56. Instruments overview. interRAI. URL: https://www.interrai.org/instruments/ [accessed 2024-04-29]
57. Monod S, Büla C, Hongler T, Castelli T, Clivaz-Luchez P. Le Resident Assessment Instrument-Home-Care (RAI-Domicile) : ce que le médecin de premier recours doit savoir [Article in French]. Rev Med Suisse 2011 Nov 09;7:2176-2183. [doi: 10.53738/REVMED.2011.7.316.2176]
58. IBM SPSS Statistics for Windows, Version 29.0. IBM Corp. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-29 [accessed 2024-04-29]
59. Cascella M, Miceli L, Cutugno F, Di Lorenzo G, Morabito A, Oriente A, et al. A Delphi consensus approach for the management of chronic pain during and after the COVID-19 Era. Int J Environ Res Public Health 2021 Dec 19;18(24):22 [FREE Full text] [doi: 10.3390/ijerph182413372] [Medline: 34948983]
60. Lam K, Iqbal FM, Purkayastha S, Kinross JM. Investigating the ethical and data governance issues of artificial intelligence in surgery: protocol for a Delphi study. JMIR Res Protoc 2021 Feb 22;10(2):e26552 [FREE Full text] [doi: 10.2196/26552] [Medline: 33616543]
61. Vogel C, Zwolinsky S, Griffiths C, Hobbs M, Henderson E, Wilkins E. A Delphi study to build consensus on the definition and use of big data in obesity research. Int J Obes (Lond) 2019 Dec;43(12):2573-2586 [FREE Full text] [doi: 10.1038/s41366-018-0313-9] [Medline: 30655580]
62. Coryell W. Depressive disorders. MSD Manual Professional Version. URL: https://www.msdmanuals.com/professional/psychiatric-disorders/mood-disorders/depressive-disorders [accessed 2024-08-29]
63. de Fruyt J, Sabbe B, Demyttenaere K. Anhedonia in depressive disorder: a narrative review. Psychopathology 2020 Jul 15;53(5-6):274-281. [doi: 10.1159/000508773] [Medline: 32668436]
64. Serretti A. Anhedonia and depressive disorders. Clin Psychopharmacol Neurosci 2023 Aug 31;21(3):401-409 [FREE Full text] [doi: 10.9758/cpn.23.1086] [Medline: 37424409]
65. Watkins ER, Roberts H. Reflecting on rumination: consequences, causes, mechanisms and treatment of rumination. Behav Res Ther 2020 Apr;127:103573. [doi: 10.1016/j.brat.2020.103573] [Medline: 32087393]
66. Wakefield JC, Schmitz MF. Feelings of worthlessness during a single complicated major depressive episode predict postremission suicide attempt. Acta Psychiatr Scand 2016 Apr;133(4):257-265. [doi: 10.1111/acps.12521] [Medline: 26538107]

67. Harrison P, Lawrence AJ, Wang S, Liu S, Xie G, Yang X, et al. The psychopathology of worthlessness in depression. Front Psychiatry 2022;13:818542 [FREE Full text] [doi: 10.3389/fpsyt.2022.818542] [Medline: 35664464]
68. Long KN, Kim ES, Chen Y, Wilson MF, Worthington Jr EL, VanderWeele TJ. The role of Hope in subsequent health and well-being for older adults: an outcome-wide longitudinal approach. Glob Epidemiol 2020 Nov;2:100018. [doi: 10.1016/j.gloepi.2020.100018]
69. Saunders R, Cohen ZD, Ambler G, DeRubeis RJ, Wiles N, Kessler D, et al. A patient stratification approach to identifying the likelihood of continued chronic depression and relapse following treatment for depression. J Pers Med 2021 Dec 04;11(12):1295 [FREE Full text] [doi: 10.3390/jpm11121295] [Medline: 34945767]
70. Haehner P, Würtz F, Kritzler S, Kunna M, Luhmann M, Woud ML. The relationship between the perception of major life events and depression: a systematic scoping review and meta-analysis. J Affect Disord 2024 Mar 15;349:145-157. [doi: 10.1016/j.jad.2024.01.042] [Medline: 38190863]
71. Voyer P. Soins infirmiers aux aînés en perte d'autonomie, 3e édition. Montréal, QC: Éditions du Renouveau Pédagogique Inc; 2021.
72. Jopp DS, Boerner K, Cimarolli V, Hicks S, Mirpuri S, Paggi M, et al. Challenges experienced at age 100: findings from the Fordham centenarian study. J Aging Soc Policy 2016;28(3):187-207. [doi: 10.1080/08959420.2016.1163652] [Medline: 27010530]
73. Schladitz K, Löbner M, Stein J, Weyerer S, Werle J, Wagner M, et al. Grief and loss in old age: exploration of the association between grief and depression. J Affect Disord 2021 Mar 15;283:285-292. [doi: 10.1016/j.jad.2021.02.008] [Medline: 33578340]
74. Kitzmüller G, Clancy A, Vaismoradi M, Wegener C, Bondas T. "Trapped in an empty waiting room"-the existential human core of loneliness in old age: a meta-synthesis. Qual Health Res 2018 Jan;28(2):213-230. [doi: 10.1177/1049732317735079] [Medline: 29235943]
75. Hernandez SC, Overholser JC. A systematic review of interventions for hope/hopelessness in older adults. Clin Gerontol 2021;44(2):97-111. [doi: 10.1080/07317115.2019.1711281] [Medline: 31913808]
76. Hafford-Letchfield T, Hanna JR, Ellmers TJ, Rasmussen S, Cogan N, Gleeson H, et al. Talking really does matter: lay perspectives from older people on talking about suicide in later life. Front Psychol 2022;13:1009503 [FREE Full text] [doi: 10.3389/fpsyg.2022.1009503] [Medline: 36467190]
77. Jopp D, Rott C. Adaptation in very old age: exploring the role of resources, beliefs, and attitudes for centenarians' happiness. Psychol Aging 2006 Jun;21(2):266-280. [doi: 10.1037/0882-7974.21.2.266] [Medline: 16768574]
78. Rott C, Jopp D. Subjective well-being in centenarians. Glob Ageing 2006;4:52-62. [doi: 10.1007/978-981-287-080-3_187-1]
79. van Wijngaarden E, Leget C, Goossensen A. Ready to give up on life: the lived experience of elderly people who feel life is completed and no longer worth living. Soc Sci Med 2015 Aug;138:257-264 [FREE Full text] [doi: 10.1016/j.socscimed.2015.05.015] [Medline: 25982088]
80. Zhang L, Li JL, Guo LL, Xu G, Yang L, Wang C, et al. Weight changes and the incidence of depressive symptom in the middle-aged and older adults: findings from the Chinese nationwide cohort study. BMC Public Health 2022 Dec 06;22(1):2282. [doi: 10.1186/s12889-022-14624-5] [Medline: 36474202]
81. Weisenbach SL, Boore LA, Kales HC. Depression and cognitive impairment in older adults. Curr Psychiatry Rep 2012 Aug;14(4):280-288. [doi: 10.1007/s11920-012-0278-7] [Medline: 22622478]
82. Muhammad T, Meher T. Association of late-life depression with cognitive impairment: evidence from a cross-sectional study among older adults in India. BMC Geriatr 2021 Jun 15;21(1):364 [FREE Full text] [doi: 10.1186/s12877-021-02314-7] [Medline: 34130632]
83. Kelly ME, Duff H, Kelly S, McHugh Power JE, Brennan S, Lawlor BA, et al. The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. Syst Rev 2017 Dec 19;6(1):259 [FREE Full text] [doi: 10.1186/s13643-017-0632-2] [Medline: 29258596]
84. Dawson-Townsend K. Social participation patterns and their associations with health and well-being for older adults. SSM Popul Health 2019 Aug;8:100424 [FREE Full text] [doi: 10.1016/j.ssmph.2019.100424] [Medline: 31312713]
85. Kandler C, Kornadt AE, Hagemeyer B, Neyer FJ. Patterns and sources of personality development in old age. J Pers Soc Psychol 2015 Jul;109(1):175-191. [doi: 10.1037/pspp0000028] [Medline: 25545842]
86. Dening KH. Differentiating between dementia, delirium and depression in older people. Nurs Stand 2019 Aug 31;35(1):43-50. [doi: 10.7748/ns.2019.e11361] [Medline: 31762251]
87. Jopp DS, Park MS, Lehrfeld J, Paggi ME. Physical, cognitive, social and mental health in near-centenarians and centenarians living in New York City: findings from the Fordham Centenarian Study. BMC Geriatr 2016 Jan 05;16:1 [FREE Full text] [doi: 10.1186/s12877-015-0167-0] [Medline: 26729190]
88. Ribeiro O, Duarte N, Teixeira L, Paúl C. Frailty and depression in centenarians. Int Psychogeriatr 2018 Jan;30(1):115-124 [FREE Full text] [doi: 10.1017/S1041610217001910] [Medline: 28931449]
89. Sinoff G. Thanatophobia (death anxiety) in the elderly: the problem of the child's inability to assess their own parent's death anxiety state. Front Med (Lausanne) 2017;4:11 [FREE Full text] [doi: 10.3389/fmed.2017.00011] [Medline: 28289681]

90. Bierman A, Upenieks L, Lee Y, Harmon M. Consequences of financial strain for psychological distress among older adults: examining the explanatory role of multiple components of the self-concept. Socius 2023 Oct 13;9(12):e1001940 [FREE Full text] [doi: 10.1177/23780231231197034] [Medline: 26701262]
91. Zaccaria D, Cavalli S, Masotti B, Gomes Da Rocha C, von Gunten A, Jopp DS. Social isolation and loneliness among near-centenarians and centenarians: results from the Fordham centenarian study. Int J Environ Res Public Health 2022 May 13;19(10):5940 [FREE Full text] [doi: 10.3390/ijerph19105940] [Medline: 35627476]
92. Cimarolli VR, Jopp DS, Boerner K, Minahan J. Depressive symptoms in the oldest-old: the role of sensory impairments. Arch Gerontol Geriatr 2018;78:249-254. [doi: 10.1016/j.archger.2018.07.009] [Medline: 30032074]
93. Li R, Zong ZY, Gu XX, Wang DN, Dong C, Sun C, et al. Higher dietary diversity as a protective factor against depression among older adults in China: a cross-sectional study. Ann Palliat Med 2022 Apr;11(4):1278-1289 [FREE Full text] [doi: 10.21037/apm-21-1449] [Medline: 34775772]
94. Madrigal-Leer F, Martìnez-Montandòn A, Solìs-Umaña M, Helo-Guzmàn F, Alfaro-Salas K, Barrientos-Calvo I, et al. Clinical, functional, mental and social profile of the Nicoya Peninsula centenarians, Costa Rica, 2017. Aging Clin Exp Res 2020 Feb;32(2):313-321. [doi: 10.1007/s40520-019-01176-9] [Medline: 30919261]
95. Merino MD, Sánchez-Ortega M, Elvira-Flores E, Mateo-Rodríguez I. Centenary personality: are there psychological resources that distinguish centenarians? J Happiness Stud 2023 Nov 08;24(8):2723-2745. [doi: 10.1007/s10902-023-00700-z]
96. Nübel J, Guhn A, Müllender S, Le HD, Cohrdes C, Köhler S. Persistent depressive disorder across the adult lifespan: results from clinical and population-based surveys in Germany. BMC Psychiatry 2020 Feb 10;20(1):58 [FREE Full text] [doi: 10.1186/s12888-020-2460-5] [Medline: 32041560]
97. Comijs HC, Nieuwesteeg J, Kok R, van Marwijk HW, van der Mast RC, Naarding P, et al. The two-year course of late-life depression; results from the Netherlands study of depression in older persons. BMC Psychiatry 2015 Feb 12;15:20 [FREE Full text] [doi: 10.1186/s12888-015-0401-5] [Medline: 25775143]
98. Pinho LG, Lopes MJ, Correia T, Sampaio F, Arco HR, Mendes A, et al. Patient-centered care for patients with depression or anxiety disorder: an integrative review. J Pers Med 2021 Aug 10;11(8):e510 [FREE Full text] [doi: 10.3390/jpm11080776] [Medline: 34442420]

Abbreviations

LLD: late-life depression

Edited by A Garcia; submitted 08.09.24; peer-reviewed by K Mikkonen; comments to author 24.01.25; revised version received 29.01.25; accepted 01.02.25; published 05.03.25.

Please cite as:
Gomes da Rocha C, von Gunten A, Vandel P, Jopp DS, Ribeiro O, Verloo H
Building Consensus on the Relevant Criteria to Screen for Depressive Symptoms Among Near-Centenarians and Centenarians: Modified e-Delphi Study
JMIR Aging 2025;8:e64352
URL: https://aging.jmir.org/2025/1/e64352
doi:10.2196/64352
PMID:40053803

Original Paper

Cyberchondria in Older Adults and Its Relationship With Cognitive Fusion, Health-Related Quality of Life, and Mental Well-Being: Mediation Analysis

Richard Huan Xu[1], PhD; Vladan Starcevic[2], Prof Dr Med

[1]Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Kowloon, China (Hong Kong)
[2]Faculty of Medicine and Health, Sydney Medical School, Nepean Clinical School, Brain and Mind Centre, University of Sydney, Sydney, Australia

Corresponding Author:
Richard Huan Xu, PhD
Department of Rehabilitation Sciences
Faculty of Health and Social Sciences
Hong Kong Polytechnic University
9 Hoi Ting Rd
Hung Hom
Kowloon
China (Hong Kong)
Phone: 852 27664199
Email: richard.xu@polyu.edu.hk

Abstract

Background: Cyberchondria is the compulsive searching for health information online that continues despite harmful effects. It leads to increased health anxiety and lower health-related quality of life (HRQOL). Older adults face higher risks of cyberchondria due to their limited digital literacy skills and more frequent health concerns. However, researchers have not thoroughly studied how cyberchondria affects this age group.

Objective: This study aimed to explore cyberchondria in the older population and investigate its relationship with cognitive fusion (ie, the tendency to become entangled with thoughts and perceive them as literal truths that dictate behavior), HRQOL, and mental well-being.

Methods: A web-based, cross-sectional survey was conducted in May 2024 with a sample of 638 participants from China aged ≥60 years recruited through the online panel of a survey company. The participants completed questionnaires assessing cyberchondria (using the Cyberchondria Severity Scale-12 [CSS-12]), cognitive fusion, HRQOL, and mental well-being. Structural equation modeling (SEM) was used to assess the hypothesized mediation model, and standardized estimates and their 95% CIs were calculated for all structural paths.

Results: Participants had a mean CSS-12 score of 40 (SD 8.5), suggesting a fairly high level of cyberchondria in this sample. Participants with a higher socioeconomic status tended to report lower levels of cyberchondria. The SEM showed that cyberchondria was positively associated with cognitive fusion (β=0.505, *P*<.001 for both models) and negatively associated with HRQOL (β=–0.221, *P*<.001) and mental well-being (β=–0.212, *P*<.001). The mediation model showed a good fit and demonstrated that cognitive fusion fully mediated the total effect of cyberchondria on HRQOL and mental well-being.

Conclusions: Cyberchondria may be more prominent in older Chinese adults, especially those residing in rural areas and with a lower socioeconomic status. Additionally, cyberchondria can enhance cognitive fusion, contributing to poor HRQOL and mental well-being. Interventions focused on “defusing” cyberchondria-relevant thoughts may help reduce maladaptive behaviors associated with cyberchondria and improve the overall well-being of older populations.

(JMIR Aging 2025;8:e70302) doi:10.2196/70302

KEYWORDS

cyberchondria; cognitive fusion; health-related quality of life; well-being; mediation

Introduction

Background

The internet can amplify health anxiety by creating an environment in which distinguishing credible online health information from that derived from unreliable sources may be difficult, often leading to confusion and fear [1,2]. Cyberchondria refers to excessive or repeated online health information seeking that persists despite negative consequences and is associated with increased health anxiety [3] and decreased health-related quality of life (HRQOL) [4,5], with its toll including diminished well-being and compromised daily functioning. Cyberchondria and health anxiety are closely intertwined, and studying this link is crucial to address the growing impact of cyberchondria. The rise of cyberchondria is intimately linked to the digital revolution, and online misinformation and worst-case scenarios often lead to catastrophic interpretations that exacerbate anxiety. The internet also facilitates confirmation bias, allowing individuals to seek information that aligns with their existing fears, reinforcing anxiety. Easy access to information facilitates compulsive checking behaviors, which provide temporary relief but ultimately perpetuate anxiety. This creates a vicious cycle, in which increased anxiety leads to more compulsive online health information seeking [4,6,7].

Cyberchondria in Older Adults

Older adults face heightened vulnerability to cyberchondria. Their vulnerability stems not only from lower digital literacy and more frequent health concerns [8,9] but also from their generation's unique beliefs and life experiences [10]. Older adults, particularly those from generations that emphasized deference to authority figures, may uncritically accept health-related information presented in an authoritative tone, even from unreliable online sources [11]. This attitude, coupled with passive health information-seeking habits developed in an era of limited access to medical knowledge, heightens susceptibility to misinformation and catastrophic health interpretations.

Moreover, older age brings an expectation of decline, a psychological state shaped by imminent health risks, accumulated losses, and direct experiences with chronic illness [12]. These experiences may foster hypervigilance toward bodily changes, prompting compulsive online searches that reinforce anxiety through confirmation bias. The interplay of factors, such as cohort beliefs, anticipatory health anxiety, and cumulative adversity, may create a unique pathway for cyberchondria in older adults, distinct from younger populations. Despite these risks, limited research has focused on cyberchondria in this age group, even as studies highlight their frequent use of the internet for health information, which may increase their anxiety and distress [12-14].

HRQOL and Cyberchondria

HRQOL is a multidimensional construct that encompasses an individual's physical health, mental health, social functioning, and role functioning [15]. HRQOL is particularly affected by cyberchondria, as the anxiety and stress from constant health-related searches have a negative impact on emotional well-being. This anxiety can also manifest as psychosomatic symptoms, affecting physical health and daily functioning.

Unravelling the mechanisms linking cyberchondria to HRQOL is essential, as individual differences in coping strategies, digital literacy, and pre-existing health conditions interact dynamically with the aforementioned specific vulnerabilities in older adults. As cyberchondria becomes more prevalent [16], targeted interventions must address not only the skill gaps but also the ingrained cognitive and emotional patterns that predispose older adults to its harms, ultimately safeguarding their well-being in an increasingly digital health landscape.

Mental Well-Being and Cyberchondria

Mental well-being and HRQOL are distinct concepts that address different dimensions of human experience. Unlike HRQOL, which quantifies the impact of health on physical, social, and role functioning, mental well-being focuses on emotional and psychological states and encompasses positive aspects such as happiness, life satisfaction, resilience, and a sense of purpose [17]. Rooted in frameworks like positive psychology, mental well-being relates to the presence of flourishing mental states (eg, optimism, self-acceptance) and absence of psychological distress (eg, anxiety, depression). Cyberchondria can jeopardize mental well-being by heightening anxiety about one's health [18]. It can also destabilize emotional health, trigger obsessive thought patterns, and impair resilience, ultimately eroding the core components of mental well-being and leaving individuals trapped in a cycle of distress and hypervigilance [19].

Elucidating the mechanisms connecting cyberchondria to mental well-being is vital, especially for older adults who face unique risks due to diminished digital literacy, heightened health concerns, and reduced adaptability to stress [20]. Factors such as pre-existing anxiety, limited ability to critically evaluate online content, and a propensity for catastrophic thinking can intensify cyberchondria's impact, making older adults particularly susceptible to its psychological toll [21]. With the growing prevalence of digital health-seeking behaviors, research must prioritize these interactions to inform robust, tailored interventions, combining enhanced digital education with cognitive-behavioral strategies, to break the cycle of anxiety and safeguard mental well-being in this vulnerable population [22].

Cognitive Fusion and Cyberchondria

Cognitive fusion refers to the tendency for behavior to be overly regulated and influenced by cognition. Cognitive fusion occurs when individuals get overly entangled with their thoughts, treating them as literal truths rather than just mental phenomena [23]. Although cognitive fusion is primarily a state phenomenon, some individuals may experience it more persistently, leading to patterns that resemble a trait. This means that certain people might be more prone to cognitive fusion across various situations, making it appear as a stable, trait-like characteristic [24]. Studies have shown that cognitive fusion is strongly associated with health care avoidance and weight stigma [25] and that higher levels of cognitive fusion are associated with

increased anxiety and depression and poorer HRQOL [26,27]. Cognitive fusion is also related to rumination, shame, and reduced self-compassion [28]. Cognitive fusion involving anxious thoughts can exacerbate anxiety-related maladaptive behaviors, such as avoidance, checking, and reassurance seeking. The inability to distance oneself from such thoughts perpetuates the anxiety cycle and strengthens the perceived validity of the thoughts [29].

Previous studies have shown an association between cyberchondria and metacognitive beliefs [30,31], a concept potentially linked to cognitive fusion. Although metacognitive beliefs refer to ways of appraising one's thoughts and other cognitive functions, cognitive fusion is about a strong attachment to one's thoughts and their perceived truth. Compared with cognitive fusion, metacognitive beliefs reflect a broader construct, can be either positive or negative, and lack an element of attachment to one's own thoughts, regardless of the nature of these thoughts. Both constructs influence how individuals experience and respond to their thoughts, but considering the potential role that cognitive fusion can play in psychopathology, it has not received sufficient attention from researchers in the context of cyberchondria.

Although cognitive fusion can affect people of all ages, older adults may be more severely impacted than their younger counterparts, given their thoughts about numerous life transitions, including retirement, loss of loved ones, high risks of social isolation, and declining health [32]. Cognitive fusion may amplify health-related anxiety of older adults by making it difficult for them to separate their concerning health-related thoughts from the reality about their health. However, limited empirical evidence exists on the relationship between cognitive fusion and health anxiety among older adults [33]. Studying cognitive fusion in this demographic may help develop tailored interventions to enhance coping mechanisms of older adults and improve their mental health and overall HRQOL [34].

The relationship between cognitive fusion and cyberchondria can be understood through the lens of cognitive processes and reinforcement. Cognitive fusion may contribute to negative perceptions of online health information [35]. In other words, the threatening thoughts about online health information, especially if this information is ambiguous or incongruent, are experienced as real and not a product of one's own interpretation or perception. This increases anxiety and makes a person search compulsively for health information in an effort to alleviate the sense of threat, thus leading to cyberchondria [29,36]. However, such behavior only reinforces the sense of threat, especially in the long run. For older adults whose health-related searches are often driven by anticipatory anxiety [12], seeking ambiguous or alarming information may deepen cognitive fusion ("I must be ill because I keep reading about symptoms"), creating a feedback loop. This aligns with studies showing that repetitive behaviors amplify cognitive rigidity in aging populations [33]. Thus, we hypothesized cyberchondria as a behavioral trigger that intensifies cognitive fusion, which in turn erodes HRQOL.

Research Questions and Hypotheses

Research Questions

Drawing from the theoretical and empirical considerations outlined earlier, this study proposed 2 research questions to guide the investigation. First, what are the levels of cyberchondria and cognitive fusion among older adults, and which sociodemographic factors correlate with these constructs? Second, how are cyberchondria, cognitive fusion, HRQOL, and mental well-being interrelated in older adults? To address the second question, 5 hypotheses were formulated. The conceptual framework is presented in Multimedia Appendix 1.

Hypothesis 1

Prior research has demonstrated that cyberchondria contributes to functional impairment and diminished quality of life [18]. In older adults, this effect may be amplified due to age-related vulnerabilities, such as reduced online health literacy and the presence of chronic illnesses, which heighten susceptibility to health anxiety and its consequences [37]. Therefore, we proposed, as hypothesis 1 (H1), that cyberchondria negatively impacts HRQOL in older adults.

Hypothesis 2

Studies have shown that excessive online health information seeking is linked to increased psychological distress, including heightened anxiety and reduced emotional well-being [38,39]. Older adults may be particularly vulnerable to these effects due to declining physical health, reduced resilience to stress, and limited digital literacy, which can exacerbate the psychological toll of cyberchondria. Thus, we hypothesized, as hypothesis 2 (H2), that cyberchondria negatively impacts mental well-being in older adults.

Hypothesis 3

Research has indicated that cognitive fusion, characterized by entanglement with thoughts and treating them as literal truths, is associated with anxiety-driven behaviors [40]. Cyberchondria, as a behavioral pattern driven by health anxiety, may similarly foster cognitive fusion by reinforcing rigid and catastrophic thought patterns about health. Accordingly, we proposed, as hypothesis 3 (H3), that cyberchondria positively impacts cognitive fusion in older adults.

Hypothesis 4

The theoretical framework of acceptance and commitment therapy, supported by studies, suggests that cognitive fusion amplifies the negative effects of maladaptive thoughts on well-being, including health-related outcomes [41]. In the context of cyberchondria, older adults who excessively seek online health information may develop rigid, catastrophic cognitions that impair HRQOL through cognitive fusion [42]. Therefore, we hypothesized, as hypothesis 4 (H4), that cognitive fusion mediates the negative impact of cyberchondria on HRQOL in older adults.

Hypothesis 5

Evidence from prior research indicates that cognitive fusion intensifies the emotional impact of anxiety-driven behaviors, mediating the association between repetitive thought patterns

and psychological distress [40]. For older adults, cyberchondria may similarly exacerbate emotional distress through cognitive fusion, undermining mental well-being [23]. Thus, we proposed, as hypothesis 5 (H5), that cognitive fusion mediates the negative impact of cyberchondria on mental well-being in older adults.

Methods

Data and Participants

This study used data from a web-based cross-sectional survey conducted between April 2024 and May 2024 to examine the health and social status of China's older population. The participants were recruited through Wenjuanxing, a Chinese survey company with an online panel of over 2.6 million members. The inclusion criteria were as follows: (1) age ≥60 years, (2) ability to read and speak Mandarin, (3) absence of cognitive impairments, and (4) ability to provide informed consent. Eligible participants were invited to complete a series of questionnaires starting with an informed consent form. Only those who consented to participate then completed the questionnaire. Several measures were implemented to ensure data quality. First, we used CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) to prevent bot submissions, ensuring that only genuine human respondents participated. We also conducted a time analysis to exclude implausibly rapid responses, which could indicate a lack of thoroughness. Additionally, we limited submissions to one per IP address within a set time frame to prevent duplicate entries. To further enhance data integrity, we identified and filtered parallel response patterns, including both consistent and repetitive patterns, which could suggest inattentive or automated responses.

Measures

The standard and validated Chinese versions of all instruments were administered. The English versions are included in Multimedia Appendix 2 for reference only.

Cyberchondria

The severity of cyberchondria was measured using the Cyberchondria Severity Scale-12 (CSS-12). It possesses good psychometric properties, comparable to those of the original version, and has been validated in the Chinese population [43]. The CSS-12 items were rated on a Likert-type scale ranging from 1 (never) to 5 (always). The total score ranges from 12 to 60, with higher scores indicating higher levels of cyberchondria.

Cognitive Fusion

The Cognitive Fusion Questionnaire (CFQ) was designed to measure the extent to which individuals are entangled in their thoughts [23]. It consists of 7 items, each rated on a 7-point Likert scale ranging from 1 (never true) to 7 (always true). Total scores range from 7 to 49, with higher scores reflecting a higher level of cognitive fusion. The psychometric properties of the CFQ have been reported to be satisfactory in the Chinese population [44].

HRQOL

The EQ-5D-5L was used to measure HRQOL in this study. It comprises 5 health-related items (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression), each rated on a 5-point Likert scale ranging from 1 (no problems) to 5 (extreme problems) [45]. To reflect HRQOL, all health states described by the descriptive system can be converted into a single utility score using a scoring algorithm based on public preferences. This study used the EQ-5D-5L China value set and scoring algorithm [46].

Mental Well-Being

The World Health Organization-5 Well-Being Index (WHO-5) is a widely recognized and validated tool for assessing subjective psychological well-being [47]. It consists of 5 questions that measure positive mood, vitality, and general interest over the past 2 weeks. Each item is rated on a 6-point Likert scale ranging from 0 (not present) to 5 (constantly present), allowing for a maximum score of 25. Higher scores indicate better well-being. The psychometric properties of the WHO-5 in the Chinese population have been reported to be satisfactory [48].

Statistical Analysis

Descriptive statistics were used to analyze participants' sociodemographic characteristics. Continuous variables (eg, age) were analyzed using means and SDs, whereas categorical variables (eg, sex) were analyzed using frequencies and percentages. Pearson correlation coefficients (r) were calculated to determine associations between variables, with r values ≥0.3 and ≥0.5 indicating moderate and large effects, respectively [49]. ANOVA was conducted to assess the differences in CSS-12 scores across different socioeconomic groups.

We used structural equation modeling (SEM) with full-information likelihood estimation to assess the hypothesized mediation model. Latent variables for cyberchondria (CSS-12), HRQOL (EQ-5D-5L), mental well-being (WHO-5), and cognitive fusion (CFQ) were created based on the sum score of the 4 instruments. Respondents' characteristics (ie, sex, age, education level, and presence of a chronic condition) were included in the model for adjustment. To assess the direct, indirect, and total effects, we used 5000 bootstrapped samples, derived effect estimates, and bias-corrected 95% CIs [50]. We assessed model fit using a Tucker-Lewis index>0.90, comparative fit index>0.90, root mean square error of approximation<0.05, and standardized root mean square residual<0.05 [51]. We set significance at P<.05 (2-tailed) and conducted all statistical analyses using R software [52].

Ethical Considerations

This study was performed in line with the principles of the Declaration of Helsinki. The study protocol and informed consent were approved by the Institutional Review Board of Hong Kong Polytechnic University (reference ID: HSEARS20230502006). Informed consent was obtained from all the participants. All data were anonymized to protect participants' privacy and ensure confidentiality.

Results

Participants' Sociodemographic Characteristics

Table 1 presents the sociodemographic characteristics of the 638 survey participants. The majority were men, comprising 64.1% (409/637) of the sample. In terms of educational background, 51.4% (328/637) had a high school education or less, 25.9% (165/637) had attended college, and 22.7% (145/637) held a university degree. Regarding their household registration, 37.5% (239/637) were from rural areas, while 62.5% (399/637) were from urban areas. When asked about their perceived socioeconomic status, 11% (70/637) considered themselves below the local average, 79.1% (505/637) felt they were equal to it, and 9.9% (63/637) viewed themselves as above it. Moreover, 60.7% (387/637) were long-term caregivers, and 55.3% (353/537) had chronic conditions.

Table 1. Participants' demographics and socioeconomic status (n=637).

Characteristics	Results, n (%)
Sex	
Male	409 (64.1)
Female	229 (35.9)
Age (years)	
60-65	533 (81.6)
66-79	105 (18.4)
Educational level	
High school or less	328 (51.4)
College	165 (25.9)
University	145 (22.7)
Household registration	
Rural	239 (37.5)
Urban	399 (62.5)
Marital status	
Single	9 (1.4)
Married	578 (90.6)
Divorced/widow(er)	51 (8)
Perceived socioeconomic status	
Lower than local average	70 (11)
Equal to local average	505 (79.1)
Higher than local average	63 (9.9)
Long-term caregiver	
Yes	387 (60.7)
No	251 (39.3)
Chronic conditions	
Yes	353 (55.3)
No	285 (44.7)

Profiles of Cyberchondria, HRQOL, Mental Well-Being, and Cognitive Fusion and Associations Between Them

Table 2 presents the participants' scores on the 4 assessment instruments, while Table 3 shows the correlations among the instruments. The average CSS-12 (cyberchondria) score was 40 (SD 8.5) points. The mean WHO-5 (mental well-being) score was 20.4 (SD 5.0) points, the mean EQ-5D-5L (HRQOL) score was 0.8 (SD 0.19) points, and the mean CFQ (cognitive fusion) score was 36.8 (SD 12.3) points. The Cronbach α coefficients for all 4 instruments were above 0.8, indicating good internal consistency and reliability.

Table 2. Measure profiles.

Measures	Mean (SD)	Median (range)	Cronbach α
CSS-12[a]	40 (8.5)	42 (12-58)	0.87
WHO-5[b]	20.4 (5.0)	21 (5-30)	0.87
EQ-5D-5L	0.8 (0.19)	0.88 (0.08-1)	0.83
CFQ[c]	36.8 (12.3_	39 (9-63)	0.94

[a]CSS-12: Cyberchondria Severity Scale-12.

[b]WHO-5: World Health Organization-5 Well-Being Index.

[c]CFQ: Cognitive Fusion Questionnaire.

Table 3. Correlations between the measures.

Measures	CSS-12[a]	WHO-5[b]	EQ-5D-5L	CFQ[c]
CSS-12				
r	—[d]	–0.20	–0.30	0.50
P value	—	<.001	<.001	<.001
WHO-5				
r	–0.20	—	0.36	–0.42
P value	<.001	—	<.001	<.001
EQ-5D-5L				
r	–0.30	0.36	—	–0.48
P value	<.001	<.001	—	<.001
CFQ				
r	0.50	–0.42	–0.48	—
P value	<.001	<.001	<.001	—

[a]CSS-12: Cyberchondria Severity Scale-12.

[b]WHO-5: World Health Organization-5 Well-Being Index.

[c]CFQ: Cognitive Fusion Questionnaire.

[d]Not applicable.

Mental well-being (WHO-5) was positively correlated with HRQOL (EQ-5D-5L; r= 0.36, P<.001), suggesting that greater well-being aligns with better HRQOL. Conversely, mental well-being was negatively correlated with cognitive fusion (CFQ; r=–0.42, P<.001), indicating that higher well-being corresponds with reduced cognitive fusion. Similarly, HRQOL was negatively correlated with cognitive fusion (r=–0.48, P<.001), demonstrating that better HRQOL is linked to lower cognitive fusion. Cyberchondria (CSS-12) had negative associations with mental well-being (r=–0.2, P<.001) and HRQOL (r=–0.3, P<.001) and a positive association with cognitive fusion (r=0.5, P<.001). These results suggest that elevated cyberchondria is associated with poorer mental well-being, diminished HRQOL, and heightened cognitive fusion (Table 3).

Differences in Cyberchondria Across Sociodemographic Groups

Table 4 illustrates the differences in cyberchondria (CSS-12) scores across various sociodemographic groups. Although no statistically significant differences were observed between the sexes, substantial variations were evident in all other group comparisons. Notably, individuals with higher educational levels, who resided in urban areas, who were younger, who had a higher socioeconomic status, who did not need caregivers, and who had no chronic conditions had lower cyberchondria scores.

Table 4. Participants' responses on the Cyberchondria Severity Scale-12 (CSS-12) and Cognitive Fusion Questionnaire (CFQ), stratified by demographics.

Demographic characteristics	Cyberchondria (CSS-12)		Cognitive fusion (CFQ)	
	Mean (SD)	*P* value	Mean (SD)	*P* value
Sex		.36		.72
Male	39.8 (8.8)		37.0 (12.4)	
Female	40.4 (7.8)		36.6 (12.2)	
Educational level		<.001		<.001
High school	41.0 (7.8)		38.5 (12.2)	
College	39.5 (8.3)		35.8 (11.8)	
University	38.3 (9.8)		34.5 (12.7)	
Household registration		<.001		<.001
Rural	42.6 (6.4)		41.6 (10.0)	
Urban	38.5 (9.2)		34.0 (12.7)	
Age group (years)		.02		.02
60-65	38.3(9.3)		34.4(12.5)	
66-79	40.3(8.2)		37.4(12.2)	
Socioeconomic status		.01		<.001
Lower than average	42.1 (8.7)		43.4 (11.0)	
Average	39.9 (8.2)		36.3 (12.1)	
Higher than average	38.5 (9.6)		34.5 (13.6)	
Do you need a caregiver?		.001		<.001
Yes	39.1 (8.9)		35.6 (12.8)	
No	41.4 (7.5)		38.9 (11.3)	
Do you have chronic condition?		<.001		<.001
Yes	41.1 (7.7)		40.0 (11.0)	
No	38.6 (9.1)		33.0 (12.7)	

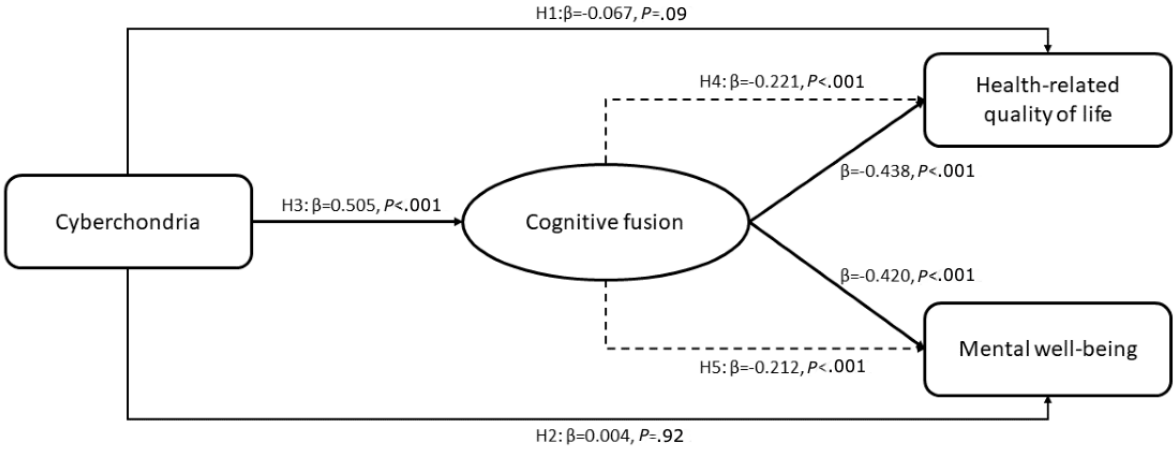
Differences in Cognitive Fusion Across Sociodemographic Groups

The differences in respondents' cognitive fusion across socioeconomic groups were similar to those found for levels of cyberchondria. Participants exhibited significantly higher cognitive fusion if they had lower educational levels, lived in rural areas, were older, had a lower socioeconomic status, and had caregivers or chronic conditions (Table 4).

Results of the Hypothesized Mediation Model

Figure 1 shows the results of the SEM analysis. The mediation model, adjusted for sex, age, and education level, demonstrated acceptable model indices (comparative fit index=0.997, Tucker-Lewis index=0.993, root mean square error of approximation=0.017), suggesting that unmeasured potential confounders were relatively unlikely to affect the results. Outcomes of the SEM analysis showed the lack of statistically significant direct effects between cyberchondria and HRQOL (β=–0.067, *P*=.09; H1 is not supported) and between cyberchondria and mental well-being (β=0.004, *P*=.92; H2 is not supported). The coefficients of the mediation paths were statistically significant. Cyberchondria was significantly positively associated with cognitive fusion (β=0.505, *P*<.001; H3 is supported) with significant negative indirect associations with HRQOL (β=–0.221, *P*<.001; H4 is supported) and mental well-being (β=–0.212, *P*<.001; H5 is supported).

Figure 1. Results of mediation analysis.



Discussion

Principal Findings

This study demonstrates that the levels of cyberchondria in the sample of older Chinese adults were rather high. Our participants had an average score of 40 (of the maximum score of 60) on the CSS-12, which surpassed the mean scores found in earlier research [53-56]. According to the CSS-12 cutoff value established by Xu [57], 60% of our participants exhibited significant features of cyberchondria. These results provide empirical evidence that internet-related mental health issues affect older adults, challenging the belief that such problems are mainly of relevance for younger individuals [58,59]. Additionally, the mediation analysis revealed that associations between cyberchondria and both HRQOL and mental well-being were complex, as no significant direct effects were observed in the model (H1 and H2 were not supported). However, cyberchondria appeared to have a negative and indirect effect on HRQOL and mental well-being by enhancing cognitive fusion among older adults (H3, H4, and H5 were supported). This highlights the importance of addressing cyberchondria not just as a mental health issue but also as a broader HRQOL concern. Interventions that aim to improve overall well-being, not just reduce anxiety, in this population should be encouraged. However, since the significant associations disappeared in the mediation models, the mechanisms relating cyberchondria to HRQOL and mental well-being might be complex and call for further research scrutiny.

The factors contributing to the high levels of cyberchondria are complex. Our study found that older adults with a lower socioeconomic status reported significantly higher cyberchondria levels than those with a higher socioeconomic status. This finding is consistent with results of previous studies conducted in other populations. For example, a Turkish study revealed a significant relationship between cyberchondria levels and family income among adolescents [60]. Several reasons may help explain this. Older individuals with a lower socioeconomic status often have limited access to health care resources, poor internet connectivity, and lower eHealth literacy [61,62]. These factors may compel them to rely heavily on potentially unreliable online health information. Moreover, older individuals with a lower socioeconomic status are more likely to experience social isolation and have smaller support networks [63,64], which may exacerbate their health anxiety and lead to higher levels of cyberchondria. Given these complexities, cyberchondria should be studied within the broader context of psychosocial determinants of health rather than as a simple health-related phenomenon.

A significant difference in CSS-12 scores emerged between urban and rural residents. Older urban residents exhibited lower levels of cyberchondria than their rural counterparts. Although a previous study found that health anxiety is more prevalent and severe in rural areas than in urban areas [65], the study empirically confirmed that the variation in cyberchondria levels was linked to household registration. The reasons for the striking difference in cyberchondria levels between urban and rural older adults largely overlap with the aforementioned reasons for the difference in cyberchondria levels between older individuals with higher and lower socioeconomic statuses. In addition, the urban migration of younger generations in China often leaves older family members in rural areas, which can hinder older people's access to direct support for using digital technologies [66]. Thus, a shift to digital health information, emphasized by the government, may become a challenge for rural residents [67,68].

The cognitive fusion scores in our sample of older adults were slightly higher than those reported in younger populations [32,69] but lower than those in populations with mental health problems, such as depression [70] and suicidal intention [71]. Gillanders et al [26] found that cognitive fusion is negatively associated with HRQOL, which is consistent with our findings. Additionally, we found a significant association between cognitive fusion and education; older respondents with higher educational levels had lower levels of cognitive fusion. Previous studies indicated that education often enhances cognitive flexibility [72], improves critical-thinking skills [73], and boosts problem-solving abilities [74]. These factors may help better-educated individuals avoid rigid thought patterns, thus reducing their proneness to cognitive fusion.

The mediation model revealed that cyberchondria may increase cognitive fusion, leading individuals to become overly entangled with their thoughts about online health information. Heightened

cognitive fusion can negatively impact HRQOL and mental well-being by making it challenging for older individuals to manage their distressing thoughts and emotions. This finding provides support to H2 of this study and partially aligns with the findings of previous studies showing that cognitive fusion significantly mediates the relationship between individual well-being and other health-related factors [75-78].

Building on previous research and clinical findings, our study extends the understanding of how cognitive fusion may serve as a potential mechanism by which cyberchondria causes negative health outcomes [79]. These findings provide valuable insights into how excessive online health information seeking may affect mental health and other functions. This study has clinical implications, as it suggests that interventions targeting cognitive fusion may be effective in mitigating the impact of cyberchondria on HRQOL and mental well-being.

As populations worldwide age, many countries are experiencing a surge in internet use among older adults, driven by increased access to smartphones and online platforms. The rising trend of cyberchondria among Chinese older adults found in this study is very relevant for other regions with aging populations, such as Japan, the United States, and much of Europe, where older adults increasingly turn to the web for health information [80]. The internet's role in amplifying anxiety through misinformation and compulsive online behavior is a universal concern [81]. Other countries can learn from China's experience to prepare for similar challenges, particularly where health care systems might struggle because of the increasing number of older patients with high levels of health anxiety and cyberchondria.

Strengths and Limitations

This is the first study conducted in China to specifically examine cyberchondria and its sociodemographic correlates in the older population. The novelty of this study lies in its examination of previously overlooked relationships between cyberchondria, cognitive fusion, HRQOL, and mental well-being in this population. The findings of this study enhance our understanding of cyberchondria by linking a specific cognitive construct (ie, cognitive fusion) with behavioral aspects (ie, excessive online health information seeking).

However, this study has several limitations that warrant consideration. First, our web-based sampling method, despite using various quality assurance techniques, may have introduced selection bias. Individuals unfamiliar with online surveys or those not part of the company's panel may have been excluded, potentially limiting the generalizability of our findings. Second, although national census data for the population aged 60 years and older was unavailable, our sample was representative only to some extent of the demographic characteristics of China's older population, as reported in previous studies. Specifically, our sample had higher proportions of educated individuals, men, and urban residents. These sampling differences may introduce bias in our findings. Third, our reliance on self-reported measures means that participants' subjective views may have influenced their responses. Fourth, the study did not assess some potentially important variables, such as digital literacy. This was due to the limited length of the study questionnaire and the need to reduce the cognitive burden on older respondents. Variables like digital literacy may affect cyberchondria, and their direct or moderating effects should be explored in future studies. Fifth, our decision to model HRQOL using the EQ-5D-5L single utility score as a manifest variable in the SEM analysis deviates from conventional SEM practices. An alternative approach, such as conceptualizing HRQOL as a latent construct with the five EQ-5D-5L dimensions as formative indicators, might better capture the multidimensional nature of HRQOL. Although this alternative does not align with the standard use of the EQ-5D-5L utility score, it might reflect the differential contributions of each dimension and adhere more closely to typical SEM frameworks. Our approach may have constrained the model's ability to account for the nuanced interplay of HRQOL dimensions, and future research should consider latent variable modeling to address this limitation. Last, our study used a cross-sectional design. Although SEM can offer some insights into potential causal relationships between variables, longitudinal data will be necessary in the future to establish definitive causal links.

Conclusion

This study shows that almost two-thirds of the older Chinese online population might exhibit features of cyberchondria. Higher levels of cyberchondria were observed among older adults living in rural areas and those with lower socioeconomic status. Targeted educational programs could be useful for empowering older adults to better navigate cyberspace and identify credible online health information. Cyberchondria can enhance cognitive fusion in older populations, resulting in poorer HRQOL and mental well-being. Interventions aimed at "defusing" cyberchondria-relevant thoughts and breaking the cycle of anxiety-fueled, excessive online health information seeking can help reduce cyberchondria and improve the overall well-being of older populations.

Data Availability

Derived data supporting the findings of this study are available from the corresponding author on request.

Authors' Contributions

Conceptualization: RHX
Data curation: RHX
Investigation: RHX
Software: RHX
Supervision: RHX, VS

Writing – original draft: RHX
Writing – review & editing:RHX, VS

Conflicts of Interest

None declared.

Multimedia Appendix 1
The conceptual framework of the study and study hypotheses. Hypothesis 1 (H1): Cyberchondria correlates negatively with HRQoL and mental well-being in older adults. Hypothesis 2 (H2): Cognitive fusion mediates the relationships between cyberchondria and reduced HRQoL and mental well-being in older adults.
[DOCX File , 30 KB - aging_v8i1e70302_app1.docx]

Multimedia Appendix 2
Survey questionnaire (English version).
[DOCX File , 27 KB - aging_v8i1e70302_app2.docx]

Multimedia Appendix 3
Factor loadings of the SEM analysis.
[DOCX File , 20 KB - aging_v8i1e70302_app3.docx]

References

1. Starcevic V. Cyberchondria: challenges of problematic online searches for health-related information. Psychother Psychosom 2017 May 11;86(3):129-133 [FREE Full text] [doi: 10.1159/000465525] [Medline: 28490037]
2. Starcevic V, Aboujaoude E. Cyberchondria, cyberbullying, cybersuicide, cybersex: "new" psychopathologies for the 21st century? World Psychiatry 2015 Feb;14(1):97-100 [FREE Full text] [doi: 10.1002/wps.20195] [Medline: 25655165]
3. Starcevic V, Berle D, Arnáez S. Recent insights into cyberchondria. Curr Psychiatry Rep 2020 Aug 27;22(11):56 [FREE Full text] [doi: 10.1007/s11920-020-01179-8] [Medline: 32852626]
4. Ambrosini F, Truzoli R, Vismara M, Vitella D, Biolcati R. The effect of cyberchondria on anxiety, depression and quality of life during COVID-19: the mediational role of obsessive-compulsive symptoms and Internet addiction. Heliyon 2022 May;8(5):e09437 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e09437] [Medline: 35600442]
5. Rahme C, Akel M, Obeid S, Hallit S. Cyberchondria severity and quality of life among Lebanese adults: the mediating role of fear of COVID-19, depression, anxiety, stress and obsessive-compulsive behavior-a structural equation model approach. BMC Psychol 2021 Oct 29;9(1):169 [FREE Full text] [doi: 10.1186/s40359-021-00674-8] [Medline: 34715930]
6. Mathes BM, Norr AM, Allan NP, Albanese BJ, Schmidt NB. Cyberchondria: overlap with health anxiety and unique relations with impairment, quality of life, and service utilization. Psychiatry Res 2018 Mar;261:204-211. [doi: 10.1016/j.psychres.2018.01.002] [Medline: 29324396]
7. Xu RH, Chen C. Moderating effect of coping strategies on the association between the infodemic-driven overuse of health care services and cyberchondria and anxiety: partial least squares structural equation modeling study. J Med Internet Res 2024 Apr 09;26:e53417 [FREE Full text] [doi: 10.2196/53417] [Medline: 38593427]
8. Ding K, Wang S, Xu W, Lin L, Liao D, Chen H, et al. Low mental health literacy and its association with depression, anxiety and poor sleep quality in Chinese elderly. Asia Pac Psychiatry 2022 Dec 09;14(4):e12520. [doi: 10.1111/appy.12520] [Medline: 36210054]
9. Chesser AK, Keene Woods N, Smothers K, Rogers N. Health literacy and older adults: a systematic review. Gerontol Geriatr Med 2016;2:2333721416630492 [FREE Full text] [doi: 10.1177/2333721416630492] [Medline: 28138488]
10. Hill S, Watts D. Googling and health anxiety in older age-a case study of CBT incorporating a single case experimental design. Clin Case Rep 2024 Aug 10;12(8):e9316 [FREE Full text] [doi: 10.1002/ccr3.9316] [Medline: 39130804]
11. Laidlaw K, Thompson LW, Dick-Siskin L, Gallagher-Thompson D. Cognitive behaviour therapy with older people. Hoboken, NJ: John Wiley & Sons Ltd; 2008.
12. Magsamen-Conrad K, Dillon JM, Billotte Verhoff C, Faulkner SL. Online health-information seeking among older populations: family influences and the role of the medical professional. Health Commun 2019 Jul 23;34(8):859-871 [FREE Full text] [doi: 10.1080/10410236.2018.1439265] [Medline: 29474125]
13. Xu RH, Tian L, Zhu L, Cao Y, Chan SK, Dong D, et al. Age differences in electronic mental health literacy: qualitative study. J Med Internet Res 2024 Jul 16;26:e59131 [FREE Full text] [doi: 10.2196/59131] [Medline: 39012686]
14. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022 Feb 16;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
15. Cheung YB, Thumboo J. Developing health-related quality-of-life instruments for use in Asia: the issues. Pharmacoeconomics 2006;24(7):643-650. [doi: 10.2165/00019053-200624070-00003] [Medline: 16802840]

16. Mestre-Bach G, Potenza MN. Cyberchondria: a growing concern during the COVID-19 pandemic and a possible addictive disorder? Curr Addict Rep 2023 Feb 09;10(1):77-96 [FREE Full text] [doi: 10.1007/s40429-022-00462-3] [Medline: 36785757]
17. Dale H, Brassington L, King K. The impact of healthy lifestyle interventions on mental health and wellbeing: a systematic review. Mental Health Review Journal 2014;19(1):1-26. [doi: 10.1108/MHRJ-05-2013-0016]
18. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother 2013 Feb 09;13(2):205-213. [doi: 10.1586/ern.12.162] [Medline: 23368807]
19. Fergus TA. Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):735-739. [doi: 10.1089/cyber.2012.0671] [Medline: 23992476]
20. Zheng H, Sin SJ, Kim HK, Theng Y. Cyberchondria: a systematic review. INTR 2020 Oct 27;31(2):677-698. [doi: 10.1108/intr-03-2020-0148]
21. White RW, Horvitz E. Cyberchondria. ACM Trans. Inf. Syst 2009 Nov 30;27(4):1-37. [doi: 10.1145/1629096.1629101]
22. Jamal A, McKenzie K, Clark M. The impact of health information technology on the quality of medical and health care: a systematic review. Health Inf Manag 2009 Oct 01;38(3):26-37. [doi: 10.1177/183335830903800305] [Medline: 19875852]
23. Gillanders DT, Bolderston H, Bond FW, Dempster M, Flaxman PE, Campbell L, et al. The development and initial validation of the cognitive fusion questionnaire. Behav Ther 2014 Jan;45(1):83-101 [FREE Full text] [doi: 10.1016/j.beth.2013.09.001] [Medline: 24411117]
24. García-Gómez M, Guerra J, López-Ramos VM, Mestre JM. Cognitive fusion mediates the relationship between dispositional mindfulness and negative affects: a study in a sample of Spanish children and adolescent school students. Int J Environ Res Public Health 2019 Nov 25;16(23):4687 [FREE Full text] [doi: 10.3390/ijerph16234687] [Medline: 31775280]
25. Harkins A, Bernent M, Hunter S. Relation between Health, Wellness, and Academic Achievement in Adolescent Females. In: J Allied Health. 2016 Presented at: ASAHP 2016 Annual Conference; October 19-21, 2016; New Orleans, LA URL: https://www.jstor.org/stable/48721952
26. Gillanders DT, Sinclair AK, MacLean M, Jardine K. Illness cognitions, cognitive fusion, avoidance and self-compassion as predictors of distress and quality of life in a heterogeneous sample of adults, after cancer. Journal of Contextual Behavioral Science 2015 Oct;4(4):300-311. [doi: 10.1016/j.jcbs.2015.07.003]
27. Guzmán A, Gillanders D, Stevenson A, Ross K. Psychosocial adjustment to mild cognitive impairment: the role of illness perceptions, cognitive fusion and cognitive impairment. Dementia (London) 2021 Feb 16;20(2):464-484. [doi: 10.1177/1471301219893862] [Medline: 31948271]
28. O'Loughlin CM, Bennett DS, O'Hayer CV. The nomological network of cognitive fusion among people living with HIV: associations with rumination, shame, and depressive symptoms. Journal of Contextual Behavioral Science 2020 Jan;15:245-252. [doi: 10.1016/j.jcbs.2020.01.012]
29. Fergus TA. I really believe I suffer from a health problem: examining an association between cognitive fusion and healthy anxiety. J Clin Psychol 2015 Sep 20;71(9):920-934. [doi: 10.1002/jclp.22194] [Medline: 25994488]
30. Fergus TA, Spada MM. Cyberchondria: examining relations with problematic internet use and metacognitive beliefs. Clin Psychol Psychother 2017 Nov;24(6):1322-1330. [doi: 10.1002/cpp.2102] [Medline: 28621035]
31. Nadeem F, Malik N, Atta M, Ullah I, Martinotti G, Pettorruso M, et al. Relationship between health-anxiety and cyberchondria: role of metacognitive beliefs. J Clin Med 2022 May 05;11(9):2590 [FREE Full text] [doi: 10.3390/jcm11092590] [Medline: 35566713]
32. Bodenlos JS, Hawes ES, Burstein SM, Arroyo KM. Association of cognitive fusion with domains of health. Journal of Contextual Behavioral Science 2020 Oct;18:9-15. [doi: 10.1016/j.jcbs.2020.08.001]
33. Sanduvete-Chaves S, Chacón-Moscoso S, Cano-García FJ. Effectiveness of psychological interventions to decrease cognitive fusion in patients with chronic pain: a systematic review and meta-analysis. J Psychosom Res 2024 Nov;186:111888 [FREE Full text] [doi: 10.1016/j.jpsychores.2024.111888] [Medline: 39213942]
34. Kállai J, Rózsa S, Hupuczi E, Hargitai R, Birkás B, Hartung I, et al. Cognitive fusion and affective isolation: blurred self-concept and empathy deficits in schizotypy. Psychiatry Res 2019 Jan;271:178-186. [doi: 10.1016/j.psychres.2018.11.045] [Medline: 30481696]
35. Barrera-Caballero S, Romero-Moreno R, Márquez-González M, Jiménez-Gonzalo L, Huertas-Domingo C, Olazarán J, et al. Longitudinal effects of cognitive fusion in depressive and anxious symptoms of family caregivers of people with dementia. Journal of Contextual Behavioral Science 2024 Jul;33:100782. [doi: 10.1016/j.jcbs.2024.100782]
36. Ford N. Modeling cognitive processes in information seeking: from Popper to Pask. J. Am. Soc. Inf. Sci 2004 Feb 23;55(9):769-782. [doi: 10.1002/asi.20021]
37. McElroy E, Shevlin M. The development and initial validation of the cyberchondria severity scale (CSS). J Anxiety Disord 2014 Mar;28(2):259-265. [doi: 10.1016/j.janxdis.2013.12.007] [Medline: 24508033]
38. Fergus TA, Russell LH. Does cyberchondria overlap with health anxiety and obsessive-compulsive symptoms? An examination of latent structure and scale interrelations. J Anxiety Disord 2016 Mar;38:88-94. [doi: 10.1016/j.janxdis.2016.01.009] [Medline: 26851749]

39. Baumgartner SE, Hartmann T. The role of health anxiety in online health information search. Cyberpsychol Behav Soc Netw 2011 Oct;14(10):613-618. [doi: 10.1089/cyber.2010.0425] [Medline: 21548797]
40. Bardeen JR, Fergus TA. The interactive effect of cognitive fusion and experiential avoidance on anxiety, depression, stress and posttraumatic stress symptoms. Journal of Contextual Behavioral Science 2016 Jan;5(1):1-6. [doi: 10.1016/j.jcbs.2016.02.002]
41. Hayes SC, Levin ME, Plumb-Vilardaga J, Villatte JL, Pistorello J. Acceptance and commitment therapy and contextual behavioral science: examining the progress of a distinctive model of behavioral and cognitive therapy. Behav Ther 2013 Jun;44(2):180-198 [FREE Full text] [doi: 10.1016/j.beth.2009.08.002] [Medline: 23611068]
42. Kashdan TB, Rottenberg J. Psychological flexibility as a fundamental aspect of health. Clin Psychol Rev 2010 Nov;30(7):865-878 [FREE Full text] [doi: 10.1016/j.cpr.2010.03.001] [Medline: 21151705]
43. Wang D, Sun L, Shao Y, Zhang X, Maguire P, Hu Y. Research and evaluation of a cyberchondria severity scale in a Chinese context. PRBM 2023 Nov;Volume 16:4417-4429. [doi: 10.2147/prbm.s431470]
44. Zhang WC, Ji Y, Li X, Guo HN, Zhu ZH. Reliability and validity of the Chinese version of the Cognitive Fusion Questionnaire. Chinese Mental Health Journal 2014;28(1):40-44 [FREE Full text] [doi: 10.3969/j.issn.1000-6729.2014.01.007]
45. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
46. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
47. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden, 12–13 February 1998. World Health Organization. 1998. URL: https://iris.who.int/handle/10665/349766 [accessed 2025-05-14]
48. Fung S, Kong CYW, Liu Y, Huang Q, Xiong Z, Jiang Z, et al. Validity and psychometric evaluation of the Chinese version of the 5-Item WHO Well-Being Index. Front Public Health 2022 Mar 30;10:872436 [FREE Full text] [doi: 10.3389/fpubh.2022.872436] [Medline: 35433612]
49. Cohen J. A power primer. Psychological Bulletin 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
50. Bollen KA. Structural Equation Models. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. Hoboken, NJ: John Wiley & Sons, Ltd; 2005:2005.
51. Kline RB. Principles and practices of structural equation modelling. New York, NY: Guilford Press; 2015.
52. R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing. 2021. URL: https://www.R-project.org/ [accessed 2025-05-14]
53. Hallit S, Rogoza R, Abi Semaan C, Azzi V, Sawma T, Obeid S. Validation of the Arabic version of the cyberchondria severity scale 12 items (CSS-12-Ar) among a sample of Lebanese adults. BMC Psychiatry 2023 Aug 23;23(1):618. [doi: 10.1186/s12888-023-05123-x] [Medline: 37612590]
54. Robles-Mariños R, Alvarado GF, Maguiña JL, Bazo-Alvarez JC. The short-form of the Cyberchondria Severity Scale (CSS-12): adaptation and validation of the Spanish version in young Peruvian students. PLoS One 2023 Oct 5;18(10):e0292459 [FREE Full text] [doi: 10.1371/journal.pone.0292459] [Medline: 37796833]
55. Barke A, Bleichhardt G, Rief W, Doering BK. The Cyberchondria Severity Scale (CSS): German validation and development of a short form. Int J Behav Med 2016 Oct 1;23(5):595-605. [doi: 10.1007/s12529-016-9549-8] [Medline: 26931780]
56. McElroy E, Kearney M, Touhey J, Evans J, Cooke Y, Shevlin M. The CSS-12: development and validation of a short-form version of the Cyberchondria Severity Scale. Cyberpsychol Behav Soc Netw 2019 May;22(5):330-335. [doi: 10.1089/cyber.2018.0624] [Medline: 31013440]
57. Xu RH. Estimation of cut-off value for the Cyberchondria Severity Scale-12 in Chinese population. Research Square Preprint posted online on August 05, 2024. [doi: 10.21203/RS.3.RS-4721206/v1]
58. Wang Q, Zhang W, An S. A systematic review and meta-analysis of Internet-based self-help interventions for mental health among adolescents and college students. Internet Interv 2023 Dec;34:100690 [FREE Full text] [doi: 10.1016/j.invent.2023.100690] [Medline: 38023965]
59. Kwak Y, Kim H, Ahn J. Impact of internet usage time on mental health in adolescents: using the 14th Korea Youth Risk Behavior Web-Based Survey 2018. PLoS One 2022 Mar 23;17(3):e0264948 [FREE Full text] [doi: 10.1371/journal.pone.0264948] [Medline: 35320295]
60. Köse S, Murat M. Examination of the relationship between smartphone addiction and cyberchondria in adolescents. Arch Psychiatr Nurs 2021 Dec;35(6):563-570. [doi: 10.1016/j.apnu.2021.08.009] [Medline: 34861946]
61. Shi Y, Ma D, Zhang J, Chen B. In the digital age: a systematic literature review of the e-health literacy and influencing factors among Chinese older adults. Z Gesundh Wiss 2023 Jun 04;31(5):679-687 [FREE Full text] [doi: 10.1007/s10389-021-01604-z] [Medline: 34104627]
62. McMaughan DJ, Oloruntoba O, Smith ML. Socioeconomic status and access to healthcare: interrelated drivers for healthy aging. Front Public Health 2020 Jun 18;8:231 [FREE Full text] [doi: 10.3389/fpubh.2020.00231] [Medline: 32626678]
63. Golden J, Conroy RM, Lawlor BA. Social support network structure in older people: underlying dimensions and association with psychological and physical health. Psychol Health Med 2009 May 14;14(3):280-290. [doi: 10.1080/13548500902730135] [Medline: 19444706]

64. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a National Academies report. Am J Geriatr Psychiatry 2020 Dec;28(12):1233-1244 [FREE Full text] [doi: 10.1016/j.jagp.2020.08.005] [Medline: 32919873]
65. Wang M, Pan Q. The rural-urban differences and influencing factors in the anxiety symptoms of Chinese elderly people. Chinese General Practice 2021;24(31):3963-3970. [doi: 10.12114/j.issn.1007-9572.2021.00.294]
66. Fan S, Yang Y. How does internet use improve mental health among middle-aged and elderly people in rural areas in China? A quasi-natural experiment based on the China Health and Retirement Longitudinal Study (CHARLS). Int J Environ Res Public Health 2022 Oct 16;19(20):13332 [FREE Full text] [doi: 10.3390/ijerph192013332] [Medline: 36293913]
67. Zhou X, Chen L. Digital health care in China and access for older people. The Lancet Public Health 2021 Dec;6(12):e873-e874. [doi: 10.1016/s2468-2667(21)00051-7]
68. Dai B. The old age health security in rural China: where to go? Int J Equity Health 2015 Nov 04;14(1):119 [FREE Full text] [doi: 10.1186/s12939-015-0224-5] [Medline: 26530528]
69. Luque-Reca O, Gillanders D, Catala P, Peñacoba C. Psychometric properties of the Cognitive Fusion Questionnaire in females with fibromyalgia. Curr Psychol 2021 Aug 30;42(11):9485-9495. [doi: 10.1007/s12144-021-02214-4]
70. Chen N, Xi J, Fan X. Correlations among psychological resilience, cognitive fusion, and depressed emotions in patients with depression. Behav Sci (Basel) 2023 Jan 25;13(2):100 [FREE Full text] [doi: 10.3390/bs13020100] [Medline: 36829329]
71. Hu Z, Yu H, Zou J, Zhang Y, Lu Z, Hu M. Relationship among self-injury, experiential avoidance, cognitive fusion, anxiety, and depression in Chinese adolescent patients with nonsuicidal self-injury. Brain Behav 2021 Dec;11(12):e2419 [FREE Full text] [doi: 10.1002/brb3.2419] [Medline: 34816613]
72. Al-Zoubi EM. Quality of life and its relationship with cognitive flexibility among higher education students. Journal of Educational and Social Research 2020 Jul 10;10(4):156 [FREE Full text] [doi: 10.36941/jesr-2020-0075]
73. Hart C, Da Costa C, D'Souza D, Kimpton A, Ljbusic J. Exploring higher education students' critical thinking skills through content analysis. Thinking Skills and Creativity 2021 Sep;41:100877. [doi: 10.1016/j.tsc.2021.100877]
74. Kanbay Y, Okanlı A. The effect of critical thinking education on nursing students' problem-solving skills. Contemp Nurse 2017 Jun;53(3):313-321. [doi: 10.1080/10376178.2017.1339567] [Medline: 28589781]
75. Faustino B, Vasco AB, Farinha-Fernandes A, Delgado J. Psychological inflexibility as a transdiagnostic construct: relationships between cognitive fusion, psychological well-being and symptomatology. Curr Psychol 2021 Jun 04;42(8):6056-6061. [doi: 10.1007/s12144-021-01943-w]
76. Fernández-Rodríguez C, Coto-Lesmes R, Martínez-Loredo V, Cuesta-Izquierdo M. Psychological inflexibility, anxiety and depression: the moderating role of cognitive fusion, experiential avoidance and activation. Psicothema 2022 May 01;34(2):240-248. [doi: 10.7334/psicothema2021.344] [Medline: 35485537]
77. Ferreira C, Mendes A, Marta-Simões J, Trindade I. The impact of early affiliative memories on psychological health: the mediator role of external shame and cognitive fusion. Eur. psychiatr 2020 Mar 23;33(S1):S518-S518. [doi: 10.1016/j.eurpsy.2016.01.1916]
78. Soltani E, Bahrainian SA, Farhoudian A, Masjedi Arani A, Gachkar L. Effectiveness of acceptance commitment therapy in social anxiety disorder: application of a longitudinal method to evaluate the mediating role of acceptance, cognitive fusion, and values. Basic Clin Neurosci 2023 Jul 01;14(4):479-490 [FREE Full text] [doi: 10.32598/bcn.2021.2785.1] [Medline: 38050569]
79. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord 2019 Feb 15;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
80. Bansal N, Choudhary H. Growing old in the digital era: a qualitative study of internet use and outcomes among urban Indian older adults. WWOP 2023 Oct 02;28(4):317-335. [doi: 10.1108/wwop-08-2023-0032]
81. Maftei A, Holman A, Merlici I. Using fake news as means of cyber-bullying: the link with compulsive internet use and online moral disengagement. Computers in Human Behavior 2022 Feb;127:107032. [doi: 10.1016/j.chb.2021.107032]

Abbreviations

CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart
CFQ: Cognitive Fusion Questionnaire
CSS-12: Cyberchondria Severity Scale-12
H1: hypothesis 1
H2: hypothesis 2
H3: hypothesis 3
H4: hypothesis 4
H5: hypothesis 5
SEM: structural equation modeling
WHO-5: World Health Organization-5 Well-Being Index

Edited by J Wang; submitted 19.12.24; peer-reviewed by S Hill, S Bidmon; comments to author 06.02.25; revised version received 01.03.25; accepted 12.05.25; published 21.05.25.

Please cite as:
Xu RH, Starcevic V
Cyberchondria in Older Adults and Its Relationship With Cognitive Fusion, Health-Related Quality of Life, and Mental Well-Being: Mediation Analysis
JMIR Aging 2025;8:e70302
URL: https://aging.jmir.org/2025/1/e70302
doi:10.2196/70302
PMID:

Exploring Literature on Data Governance in the Health Care of Older Persons: Scoping Review

Angely Garcia[1,2,3*], MPH, RN; Ayra Mae Balingbing[1*], PTRP; Yves Palad[3*], EdD, PTRP

[1]Institute on Aging, National Institutes of Health, University of the Philippines Manila, 623 Pedro Gil St Ermita, Manila, Philippines

[2]College of Nursing, University of the Philippines Manila, Manila, Philippines

[3]College of Allied Medical Professions, University of the Philippines Manila, Manila, Philippines

[*]all authors contributed equally

Corresponding Author:
Angely Garcia, MPH, RN
Institute on Aging, National Institutes of Health, University of the Philippines Manila, 623 Pedro Gil St Ermita, Manila, Philippines

Abstract

Background: Health data are growing rapidly, and the processing of such data is evolving. Research on data governance in older persons' health care is unexplored, providing little guidance for practice and future studies.

Objective: This scoping review aimed to synthesize available information on data governance in the context of older persons' health based on evidence from literature.

Methods: The study followed the methodological framework of Arksey and O'Malley and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews). Online databases, namely, PubMed, Cochrane, Ovid, ACM, IEEE Xplore, and Google Scholar were searched by 2 independent reviewers (AG and AMB) for studies on older persons' health data governance published from January 2000 to April 2024. The independent reviewers performed the search, screening, data extraction, and review of full-text papers. A third reviewer (YP) made the final decision for unresolved discrepancies between the first 2 reviewers. The framework by the World Health Organization Pan American Health Organization, a high-level framework for planning and implementing data governance in public health, was used in the data extraction and analysis. Descriptive statistics were used, and a descriptive approach was used to summarize the results of the scoping review.

Results: A total of 9840 titles were identified and 57 papers were included. Of these, 35 (61.4%) focused on technology, 19 (33.3%) on processes, and 3 (5.3%) on people. Data controller, processor, researchers, data subject or patient (including family or relatives), and relevant organizations were involved in older persons' data governance. Data governance frameworks were designed and implemented by reviewing the current evidence, involving the stakeholders throughout the process, implementing specific procedures (eg, collection and aggregating health data), and monitoring and evaluating them.

Conclusions: The review underscores the importance of the involvement of relevant stakeholders and the use of various innovative tools and approaches in governing data related to the health of older persons. Meanwhile, research specifically addressing data governance for older persons' health conditions is limited. To enhance health outcomes for older persons, effective data governance is essential, alongside further research on relevant policies and practices.

(JMIR Aging 2025;8:e73625) doi:10.2196/73625

KEYWORDS

older persons; health; data governance; scoping review; medical records

Introduction

Data governance is a management discipline and an emerging data management approach in health care. It is defined within the context of information technology as specifying the framework for decision rights and accountabilities to vitalize desirable behavior in the use of data [1]. Data governance highlights the responsibilities of those with authority in the organization [2], internal and external stakeholders [3], such as data stewards [4], civic society, public bodies [5], professional bodies [6], and the individuals who contributed their data [4], in managing data through its life cycle to generate quality information that can inform decision-making. Data life cycle begins with the capture or collection to the processing, use, storage, and disposal of data [7].

Data governance is also important in health systems as it is considered by many organizations as a promising method of maintaining health data [8]. Health data are considered an important asset to improve health through their use in public health, epidemiology, and health informatics [9]. Governance of such data allows health organizations to successfully manage, protect, maintain, and use data to generate information that

improves health care quality, health outcomes, and health system performance [10].

Real-time generation and efficiency in obtaining knowledge are possible when data governance principles are applied in health care [3]. In contrast, the absence of data governance could lead to failure in decision-making and addressing the individual needs of the public sector [11]. Literature underscores the impact of data governance on a nation's health care system, and the need to collect the right data, effectively process it to generate quality information for evaluating the health system, and identify where and when it is not functioning well [12]. Available evidence calls for the need to establish, streamline, and institutionalize a strong and comprehensive data governance process [5,13-19].

Data governance is essential in digital transformation initiatives of organizations as it improves data quality and accuracy and facilitates real-time data exchange [20]. One of the principles of the global health sector's digital transformation is accelerating progress toward inclusive digital health with emphasis on the most vulnerable populations [15]. These vulnerable populations include the older persons. The potential of data governance in the field of older persons' care has been highlighted by the World Economic Forum by presenting a new approach to data governance [21]. It emphasized that the combination of caregiver skills, for example, with the older persons' data such as their specific care needs, can result in better and more precisely tailored care for older persons.

The state of research on data governance in the health care of older persons is unexplored locally and internationally. This leaves little guidance for its application in practice and for future research on the topic. Thus, this scoping review aimed to map out the available literature on data governance in the context of older persons' health. Specifically, this review sought to answer the following questions: Who is involved in the planning and implementation of data governance ("people")? How are data governance frameworks designed and implemented? What are the governance processes that lead to the improvement of older persons' health ("process")? What tools and technologies are used to effectively govern data ("technology")?

Methods

Study Design and Framework

This scoping review followed the methodological framework of Arksey and O'Malley [22] and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [23]. The review questions were informed by the data governance framework in public health by Pan American Health Organization (PAHO), a high-level framework for planning, implementing, and continuously improving data governance [10]. The framework highlights the involvement of 3 components: people, processes, and technology. The people component includes establishing decision-making structures (eg, both the executive decision-making and the technical bodies); defining roles and responsibilities of those involved in the management of data throughout its life cycle; communicating these roles and responsibilities, data-related decisions, policies, and processes; and ensuring transparency on adherence to standards. Processes include managing data assets, enabling processes and standard operating procedures, and establishing processes for policy management and ensuring standards. The technology component pertains to identifying and implementing tools and technology required to manage data such as hardware and software; ensuring quality, availability, and security of information systems; and ensuring performance of the tools [10].

Aside from the framework for planning and implementing data governance, the data governance functions throughout the data life cycle also guided this scoping review. Data governance should address and include defining accountabilities, prioritizing investment requirements, establishing policies, implementing processes, setting standards, managing risks, and monitoring performance related to data throughout its life cycle. The data life cycle is as follows: data collection; data aggregation; data quality assurance and monitoring; data storage; data protection; data access, use, and disclosure; and data retention and destruction [10].

Sources of Information

Six databases, namely, PubMed, Cochrane, Ovid, IEEE Xplore, Association for Computing Machinery (ACM), and Google Scholar were searched. PubMed, Ovid, and Cochrane contained a vast collection of health literature. IEEE Xplore and ACM digital libraries contained resource materials from the fields of electrical engineering, information technology, computer science, and electronics, which capture the technology-related aspect of this review. Google Scholar indexes the full text or metadata of scholarly literature and contains a wide variety of disciplines and sources.

Search Strategy

The strategy was developed in consultation with an academic librarian and the search terms used were "data governance," "health data governance," "older person," "older people," "older adult," "elderly," "senior citizen," and "aged." Boolean operators were used to filter search results (Table S1 in Multimedia Appendix 1).

This review also included evidence from nonresearch sources [24] such as reports, projects, and economy papers from Google Scholar. Moreover, experts in data governance, eHealth, and geriatrics were consulted to explore additional literature sources. The experts were identified through the existing networks of the investigators and organizations relevant to digital health. Searching the reference list of the identified data governance literature was also conducted.

All the literature searched through the online databases, digital libraries, consultation with experts, and reference listing was uploaded in a shared Google Drive folder accessible only to the reviewers using their official university email addresses. The titles were encoded using Google Sheets to identify and remove duplicates.

Selection Process

Considering that data governance was introduced in 2000 and literature on the topic started to be published around the same

time [1,2,25], the search was filtered by publication dates between January 1, 2000, and April 22, 2024. The papers that meet the following criteria were included: (1) topic related to data governance including people, processes, or technology, (2) topic relevant to older persons' health care, (3) peer-reviewed publications, reports, policies, programs, and policy briefs, and (4) in the English language. Resources that were not related to data governance and older persons' health, conference abstracts, and published in languages other than English were excluded.

Following the eligibility criteria, 2 reviewers (AG and AMB) independently searched databases, screened papers for eligibility, reviewed the full texts, and extracted data from the included studies. If the reviewers are unable to screen papers based on the title alone, abstract screening was done to check whether the papers are related to health, data governance, and older persons. Initial data extraction from the abstracts was done to document the topic they covered using a data extraction form.

The reviewers independently documented the search yields and listed the titles obtained per online database using Google Sheets. The search yields of each reviewer per online database were compared as a form of initial validation. To ensure consistency among the reviewers, the procedures from searching to data extraction were pilot tested using 10 randomly selected samples. Reviewers proceeded with the next phases only upon reaching 75% agreement [24].

For each phase, the reviewers documented the reasons for exclusion and settled discrepancies by discussing them throughout the selection process. The reviewers met online and conferred about the discrepancies, which led to an amendment of the eligibility criteria. Based on the initial assessment of the abstracts and available full texts, there were various papers with mixed populations as their study samples (eg, 18 years and above). Considering the focus of this review, the reviewers decided to only include papers with older persons explicitly stated in their titles or abstracts or representing more than 50% of the study sample. Meetings between the reviewers were conducted to discuss which of the papers would be included in the review.

Data Extraction

Once the list of included studies was finalized, the 2 reviewers (AG and AMB) independently extracted the relevant information using an extraction tool in Google Sheets. Information such as authors, citation, publication year, type of publication, research design, setting of care, population, data governance function and component, and outcomes of the study and the technological intervention, among others, were extracted for further analysis. The reviewers conducted regular meetings to discuss and reconcile differences in the extracted data. Any disagreements that remained unresolved were discussed with a third reviewer (YP) to arrive at a consensus. The third reviewer made the final decision for the unresolved disagreements between the first 2 reviewers.

Data Management and Analysis

The search results and study selection process are documented in this scoping review report and presented using a PRISMA-ScR flow diagram. All data were documented in a Google Sheets document and exported into an Excel file after the entire process ended. For the unstructured texts in the data extraction form, a qualitative content analysis was conducted to describe how data governance is planned and implemented in the care of older persons. The data were coded according to the framework components involved, namely, person, processes, and technology, and data governance functions. A descriptive approach was used to summarize the result of the scoping review. Numerical or categorical data were presented using counts and proportions. Tables and a figure were used to present the extracted data for each extraction category.

Ethical Considerations

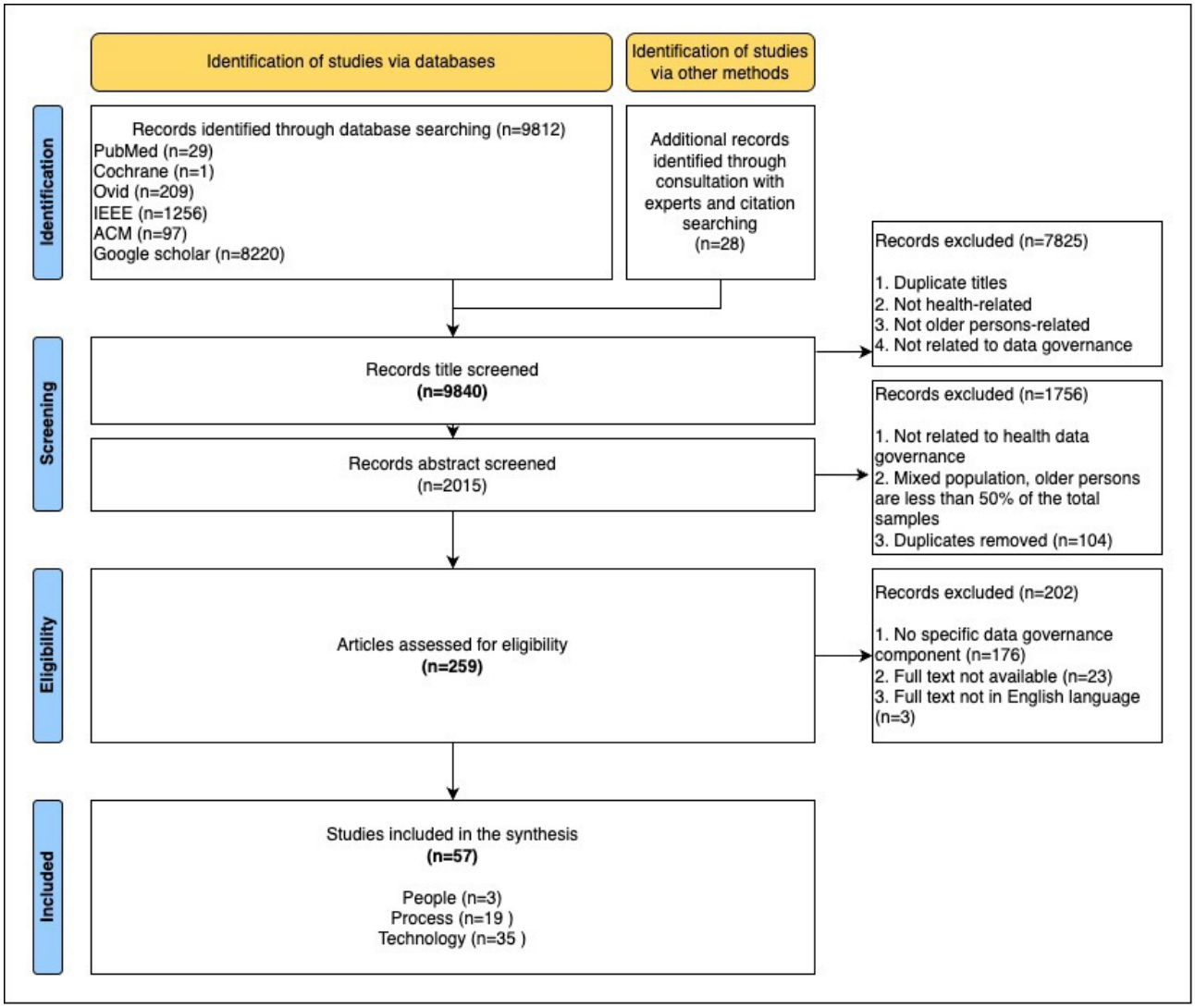
The study was monitored and classified as exempt from ethics review (UPMREB 2023 - 0864-EX) by the University of the Philippines Manila Research Ethics Board.

Results

Overview

A total of 9840 titles were identified from 6 databases, consultation with 3 experts, and reference listing (Figure 1). Duplicates and all papers not related to older persons, health, and data governance, and those with mixed populations were removed during the screening phase, leaving a total of 259 papers assessed for eligibility. Of these papers, 202 were removed due to the unavailability of full texts, non-English language, and lack of specific data governance components. A total of 57 (22%) papers met the criteria for inclusion in the synthesis.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) diagram.



Description of Included Studies

The included studies were published between 2003 and 2024; 56.1% (32/57) of these were published in the year 2020 onwards. The majority (32/57, 56.1%) of the included studies were from the IEEE digital library. Most (47/57, 82.5%) were journal papers, followed by conference proceedings (7/57, 12.3%). These conference proceedings were peer-reviewed full papers from IEEE and ACM (Table 1).

In terms of study design, 36.8% (21/57) were observational, followed by 29.8% (17/57) classified as experimental studies. These experimental studies are subdivided into 2 groups, namely, technology development and randomized controlled trials. About 30% (16/57) of the studies did not specify their design.

The studies were conducted mostly in home settings (11/57, 15.8%), followed by those conducted in multiple settings of care such as home and hospital, community and hospital, and hospital and long-term care facilities. More than 30% of the included studies did not specify their study setting. In terms of geographic coverage, most were conducted in Europe, followed by Asia and those with multicountry sites. There were no included studies that covered the following continents: Africa, Middle East, and South America (Table 1).

Sixty-seven percent of the included studies provided information about the study population, particularly, the age group, mean age, or age range they covered. Although 33.3% (19/57) did not indicate the specific age of their study population, the technology or the data governance component being relevant to older persons was explicitly stated in the title, abstract, and full text. Due to the variation of available information on age and unavailability of other information, the overall mean age cannot be computed.

About 70% of the included studies specified the sample size. A study had 1 older person sample, which was about a novel cloud-based framework for the older adult health care services [26]. The largest sample size among the included studies was 591,726 electronic health records of primary care patients aged 65 years and older [27].

Table . Profile of included studies.

Description	Included studies (n=57), n (%)
Data sources	
IEEE	32 (56.1)
Google Scholar	15 (26.3)
ACM[a]	4 (7.0)
PubMed	4 (7.0)
Ovid	2 (3.5)
Expert consultation	0 (0)
Publication year	
2000 - 2009	4 (7.0)
2010 - 2019	21 (36.8)
2020 and onward	32 (56.1)
Publication type	
Journal paper	47 (82.5)
Conference proceeding	7 (12.3)
Thesis	1 (1.8)
Review	1 (1.8)
Economy paper	1 (1.8)
Study design	
Observational	21 (36.8)
Experimental	17 (29.8)
Technology development	15 (26.3)
Randomized controlled trial	2 (3.5)
Review (eg, systematic review and scoping review)	3 (5.3)
Not specified	16 (28.1)
Study settings	
Home	11 (19.3)
Multiple settings	9 (15.8)
Hospital	5 (8.8)
Long-term care setting	5 (8.8)
Community	4 (7.0)
Others: research laboratory or special facility for experiment, university, office	4 (7.0)
Not specified	19 (33.3)
Geographic coverage	
Europe	12 (21.1)
Asia	6 (10.5)
Australia and Oceania	3 (5.3)
North America	2 (3.5)
Multicountry or continent sites	6 (10.5)
Not specified	28 (49.1)
Age (years) of older persons in study sample	

Description	Included studies (n=57), n (%)
Specified (age covered/mean/range)	38 (66.7)
Minimum age	18 (N/A[b])
Maximum age	104 (N/A)
Not specified	19 (33.3)
Sample size	
Specified	39 (68.4)
Minimum sample size	1 (N/A)
Maximum sample size (datasets)	591,726 (N/A)
Not specified	18 (31.6)

[a]ACM: Association for Computing Machinery.

[b]N/A: not applicable.

Data Governance Components and Functions

A majority (35/57, 61.4%) of the included studies were largely related to technology, tools and technology, in particular (Table 2). This is followed by processes (19/57, 33.3%), which are mostly centered on enabling processes and standard operating procedures. Only 3 papers were focused on people, particularly, roles and responsibilities and decision-making structure.

Interestingly, there were no included studies on communication and transparency under the people component, standards and policy management in the process component, availability and security, and performance alone in the technology component (Table 2).

Table 2 also shows that the data governance function of the included studies mostly focused on implementing processes, while only a few covered monitoring performances [27-30], defining accountability [31-33], setting standards [34,35], and managing risks [36].

Table . Data governance components and functions of included studies.

Data governance	Included studies (n=57), n (%)
Data governance component	
People	3 (5.3)
None	54 (94.7)
Roles and responsibilities	2 (3.5)
Decision-making structure	1 (1.8)
Communication and transparency	0 (0)
Process	19 (33.3)
None	37 (64.9)
Enabling processes and SOPs[a]	11 (19.3)
Data asset management	8 (15.8)
Standards and policy management	0 (0)
Technology	35 (61.4)
None	22 (38.6)
Tools and technology	33 (57.9)
Tools and technology and performance	2 (3.5)
Availability and security	0 (0)
Performance	0 (0)
Data governance function	
Implement processes	47 (82.5)
Monitor performance	4 (7.0)
Define accountability	3 (5.3)
Set standards	2 (3.5)
Manage risks	1 (1.8)
Prioritize investment	0 (0)
Establish policy	0 (0)

[a]SOP: standard operating procedure.

Designing and Implementing Data Governance

The identified actors involved in designing and implementing data governance related to the health of older persons include data controllers, processors, and subjects. Specific data controllers identified in this review include a hospital [32], researchers [33], and a steering committee [37]. The data subjects were patients [32], persons with dementia [33], nursing home residents [37], friends or relatives [32], legally authorized representatives [33], and hospital staff [32]. The data processors were a platform service provider [32], research team members [33], and privacy, scientific expert, and data access committees [37].

All included studies on the people component of data governance highlighted the crucial roles of relevant stakeholders (eg, controllers, processors, and subjects) in planning, designing, and implementing data governance [27,32,33]. Moreover, the composition, roles, and accountabilities of these stakeholders vary depending on the settings (eg, hospital, research, and nursing home).

An example specifying the roles and responsibilities of people involved in data governance was highlighted by Bernsmed [32] in the context of medical data collected from sensors that are exchanged between the older persons, their families and friends, and health personnel. The roles of data controller, processor, and subject were defined, and the accountability obligations were outlined. A data controller, which can be a person, public authority, agency, or other actors, determines the purposes and means of personal data processing. The data processor processes the personal data on behalf of the controller. Meanwhile, a data subject is a person who can be identified directly or indirectly through a reference identification number. In the study, the patients and relatives or friends, as well as the hospital staff, were the data subjects, while the hospital is the controller of its patients' and personnel's data. The data controller (hospital) is accountable to patients, relatives or friends, and hospital staff (data subjects) in informing and obtaining consent for the collection, processing, and management of their data [32].

A total of 10 papers provided information on how frameworks related to data governance of older persons' health data are

designed and implemented [26,32,33,35,37-42]. Specific examples of the use of data governance frameworks were related to addressing a particular issue such as fall detection [41], dementia care mapping [42], depression risk prediction [35], cardiac image processing [40], and digital twin health care for real-time supervision and crisis warning [43].

Conceptual content analysis of the methods using the PAHO framework emphasized the involvement of relevant stakeholders, implementation of specific processes, and conduct of literature review, demonstrating the feasibility or testing of the developed framework and defining the roles and accountabilities of the actors or stakeholders (Table S2 in Multimedia Appendix 1).

Data Governance Processes, Tools, and Outcomes

A total of 19 papers provided information on the specific processes or interventions relevant to older persons' health [26,30,34,35,38,39,41-53]. Majority focused on implementing processes pertaining to data collection, aggregation, and access, use, and disclosure. One study covered 6 data life cycles from data collection up to disclosure [44]. There were no papers that covered the process of defining accountability, managing risks, prioritizing investment, and establishing policy (Table 2).

Several tools and technologies were identified and used for data governance throughout the data life cycle. A total of 35 papers provided information on the specific technologies relevant to older persons' health [27-29,36,40,54-82]. Most of the identified technologies focused on implementing processes while only a few were related to use of technology in monitoring performance, defining accountability, and managing risks. Furthermore, these technologies are mostly involved in data collection.

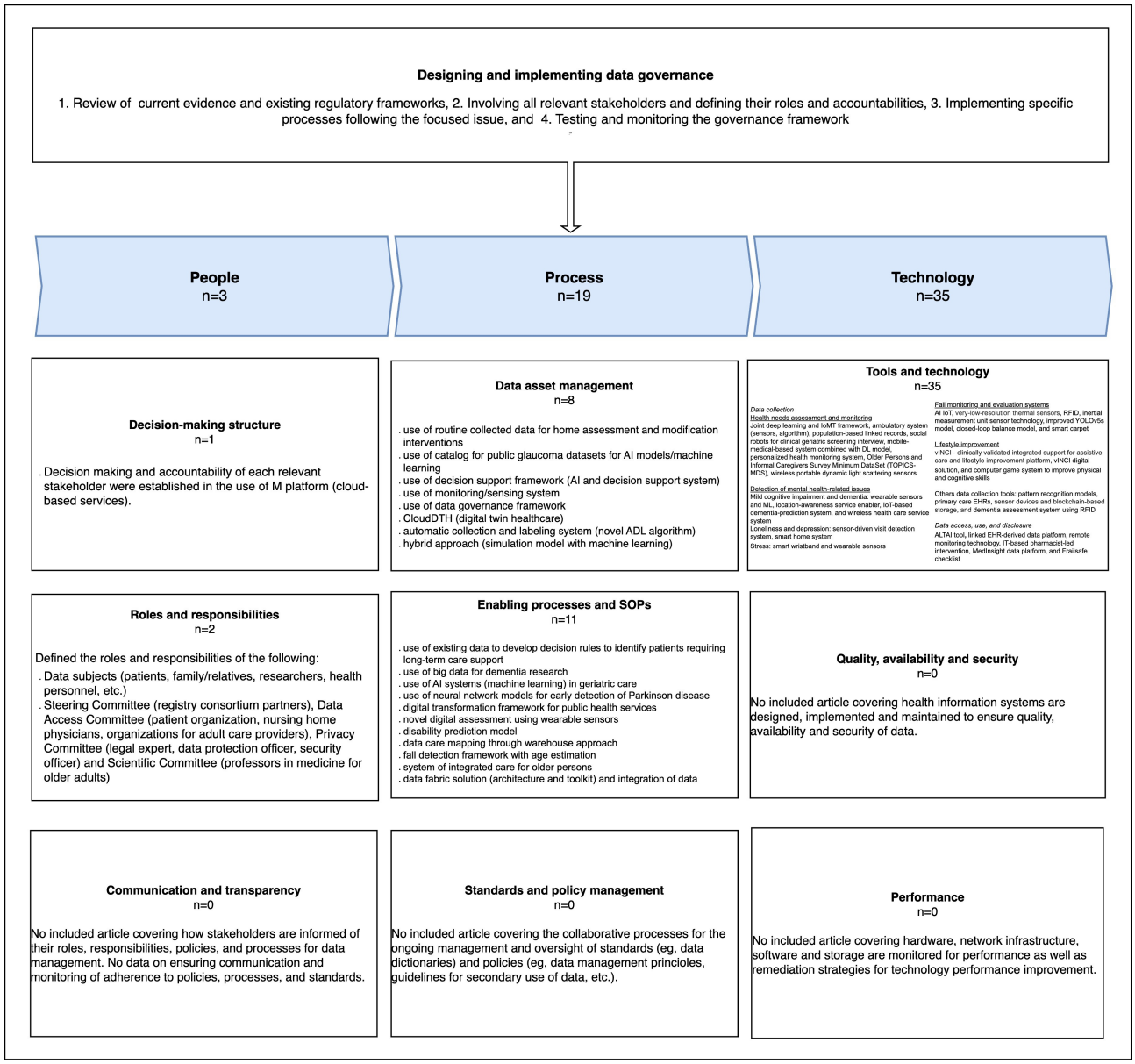
In implementing data governance processes, various tools and technologies are used in collecting data for the assessment and monitoring of health needs; assessment of fall risk and detection of fall; identification of mental health–related issues such as mild cognitive impairment, dementia, stress, loneliness, and depression; and improvement of lifestyle and health (Table S2 in Multimedia Appendix 1).

In terms of specific outcomes of the technology component of data governance, a total of 34 papers provided information on these. Most of the outcomes were process-related, such as efficient data access; quality control; communication and care support; recognition, detection, and prediction of various conditions; monitoring; and others. Only 2 included papers provided information on health outcomes among the older person samples, namely, improvement of quality of life [62] and cognitive skills and function [68] (Table S2 in Multimedia Appendix 1).

Older Persons' Health Care Data Governance

Figure 2 summarizes the results of this scoping review following the data governance framework in public health by PAHO [10].

Figure 2. Older person's health care data governance based on the available literature. ADL: activities of daily living; AI: artificial intelligence; EHR: electronic health record; IoT: Internet of Things; DL: deep learning; ML: machine learning; RFID: radio frequency identification; SOP: standard operating procedure.



Discussion

Principal Results

This scoping review aimed to synthesize available information on how data governance is planned and implemented in the context of older persons' health based on evidence from literature. Specifically, it aimed to answer the following questions: Who are involved in the planning and implementation of data governance (people)? How are data governance frameworks designed and implemented? What are the governance processes that lead to the improvement of older persons' health (process)? What tools and technologies are used to effectively govern data (technology)?

Findings show that available studies related to the health of older persons covered mostly data governance components related to technology and process. Moreover, data governance is designed and implemented by initially examining the current evidence and existing regulatory frameworks, involving all relevant stakeholders and defining their roles and accountabilities, implementing specific processes following the focused issue, and testing and monitoring the governance framework. Various data governance processes, tools, and technologies are used and contribute to the improvement of older persons' health. The use of these mostly leads to process outcomes.

This review emphasizes the involvement of stakeholders, particularly in technology development and testing process [40,41], sharing perspectives on digital transformation of public health services [39], and mapping of service users [42]. Moreover, the accountability of data controllers, processors, and subjects was outlined [32,33,37]. This underscores the data governance responsibilities of those with the authority in the organization [2] (data controller), the internal and external stakeholders (data processors and subjects) [3-6], and, ultimately, the data subjects who contributed their data [4] in

managing data through their life cycle to inform decision-making.

Limitations

This review included literature only on data governance relevant to older persons published from the past 2 decades and those peer-reviewed publications, reports, policies, programs, or policy briefs written in English language. Identified studies eligible for review but without full texts were excluded from the synthesis. Despite the efforts of the reviewers and consultant librarian in accessing the papers, 23 full texts were not retrieved. These papers were mostly from ACM and Google Scholar, published 2021 and beyond, included conference proceedings, and focused on the technology component of data governance. These exclusions might have affected the comprehensiveness of the reported findings on technology.

Many included studies did not specify the study design, setting, geographic coverage, age of the study population, and sample size. It would be helpful if this information were available to provide a complete description of the state of data governance literature for older persons. Moreover, the assessment or appraisal of the included studies for methodological rigor and quality was not performed as this review aimed to describe information only on data governance in the context of older persons' health.

Various processes and technologies aimed to contribute to the improvement of older persons' health based on their objectives and purpose were identified. However, since most of the included studies are related to technology, at their preclinical stage of technological development, and focused only on data collection, limited information on the actual health outcomes of data governance specific to older persons was collected.

Comparison With Prior Work

Studies related to health information systems included in this review [27,60,79] also support data governance as an important player in the health system, a promising method of maintaining health data [8], and recognize its potential in the field of older persons' care [83,84]. Various processes, tools, and innovative approaches in governing health data of older persons were explored in this review. Available evidence suggests that the use of these tools and technologies supports data governance, leading to positive health outcomes and better processes [27,31,61-63,68,75-77]. Although positive health outcomes were noted, further interventional studies with utilization of a data governance framework are essential to determine its direct effects on actual health outcomes among the older population.

In terms of processes, findings in this review are consistent with Cave [85] where strategies for implementing data governance were explored through a qualitative case study. These strategies include structured oversight with committees and boards, obtaining stakeholder buy-in, and benchmarking and standardization. Benchmarking and standardization through review of current evidence and existing regulatory frameworks were identified to be important steps in planning and implementing data governance [32,33,35,38,39,42]. These preliminary steps are crucial in ensuring a well-designed and effective governance framework.

Strategies for effective and strategic communications and compliance with regulations were not adequately covered in this review due to the lack of studies on this topic. Communication and transparency, as well as compliance with standards and policies, should be considered in implementing data governance. This is to ensure the protection of people, promotion of health value, and prioritization of equity, which are the key health data governance principles [86]. Moreover, although there were no papers that covered the processes of defining accountability and managing risks, these were discussed in the literature in terms of the role of persons involved in data governance and as integral components of designing and implementing data governance.

Various geriatric syndromes, such as falls and fall risk [54,55,65,69,71,73,74], mild cognitive impairment and dementia [33,42,53,64,66,75,82], frailty [29,80], loneliness, social isolation, and depression [56,57], were the focus of the studies included in this review. However, there were no included studies related to other geriatric syndromes such as functional decline, incontinence, delirium, pressure ulcers, polypharmacy, malnutrition, sleep problems, and others. These geriatric syndromes are common health conditions among older persons [87-90], often having multifactorial causes, and may have a major impact on their quality of life and disability [90]. Data governance covering these conditions can also generate information on strategies to address the needs of older persons and provide quality care to improve their health and quality of life.

Only 1 paper explicitly proposed a data governance framework specific for older persons, which was published a decade ago [38]. The proposed framework was drawn from the Data Management Body of Knowledge of the DAMA International, Inc [2,91] and the work of Cleven and Wortmann [92]. Dahlberg's [38] motive for proposing this specific governance framework was the fragmentation of older persons' data and the necessity to consolidate these data to make it more useful. The paper defined the types and sources of data that exist about older persons, and how these can be integrated into a comprehensive framework. It suggested defining the data categories, attributes, and sources of data first to improve data governance of older persons. There is no evidence that the framework has been validated. Hence, testing this framework is essential to determine its value in data governance for older persons.

It is notable that the included papers that provided information on designing and implementing data governance mostly came from high-income countries such as Australia, United States, Canada, United Kingdom, Japan, Singapore, Netherlands, and other countries in Europe. These countries are actively implementing digital health care solutions in the forms of telemedicine, electronic records, mobile health, and other digital tools for diagnosis and management, hence, the availability of data governance frameworks. On the other hand, no studies were included in the review from Africa, Middle East, and South America. This can be attributed to the challenges experienced in digital health, such as limited infrastructure, affordability issues, and data privacy concerns, especially in low-income and middle-income countries [93-95]. While others have established

and are implementing data governance, the rest are yet to address the challenges of digital health transformation.

The need for data governance in general [14] and in the care of older persons has been highlighted by the reviewed papers specifically on dementia research [42,53], health care applications for geriatric clinical care [49], and nursing homes for quality improvement [37]. New paradigms on the use of big data in dementia research and clinical care [42], as well as international dementia care mapping, require data governance [53]. Likewise, the standardization of machine learning approaches tailored to health care applications is required to evaluate whether these applications improve clinical care for older persons [49]. In nursing homes, literature suggests the need for enhancing transparency, specifically in presenting understandable information to the residents and their representatives on which data will be used, how they will be used, and for what purposes [37].

Findings agree with other papers not included in this review such as the call for institutionalizing a strong and comprehensive data governance process [5,10,12-19,96-98], a national health data governance framework to ensure availability and use of personal health data for public interest [99], and data governance for learning health systems to reduce concerns about privacy and trust in the system [100,101].The included studies were analyzed according to the data life cycle, data governance function, and components they covered to identify research gaps. The people data governance component covered only data access, use, and disclosure. Information on how people are involved, their roles and responsibilities, and communication and transparency during data collection, aggregation process, quality assurance, storage, data protection, retention, and destruction, is lacking. In terms of the process component, limited studies were related to data quality, storage, and protection, and none for defining accountability and managing risk. Meanwhile, no tools and technology were identified from the studies that covered data aggregation, data quality, data protection, and setting standards.

Studies on data governance in health care of older persons are largely focused on various forms of technology followed by process. The major research gap across all data governance components was the lack of included studies on data retention and destruction, prioritizing investments, and establishing policy (Table S3 in Multimedia Appendix 1). Findings and gaps that were identified from this review inform future research directions and practices of older persons' health data governance.

Conclusions

This review highlighted the importance of benchmarking, involvement of relevant stakeholders, and the use of various innovative tools and approaches in governing data related to the health of older persons. However, studies that are explicitly centered on data governance of older persons' health data, their common health conditions, or other geriatric syndromes, are limited. Likewise, identified research gaps include the lack of studies on data retention and destruction, establishment of policy, prioritizing investment, and communication and transparency.

Available studies on the health of older persons covered mostly technology and process data governance components. Only a few focused on people, which were identified to have crucial roles in designing governance frameworks, planning, and implementing data governance. To optimize the use of data in improving the health and quality of life of older persons, a well-designed process that considers the essential data governance components, functions, and principles should be developed and implemented.

Recommendations

Effective governance of older persons' health data requires a multisectoral approach. Benchmarking, utilization of innovative tools and approaches, and collaboration between relevant stakeholders, including policy makers, program planners, health facility administrators, health care providers, information and communication technology and digital health experts, data privacy and legal experts, and, ultimately, the older persons and their families, are essential in designing and implementing data governance.

In establishing health data governance policies or incorporating them into the national standards, legislations, and organizational policies, the key governance principles, components, functions, and data life cycle should be considered. In practice, health settings and personnel should ensure the involvement of the patients (including their families and relatives) in collecting and processing their data. The practical application of data governance frameworks and technologies in the care of older persons is essential to support favorable health outcomes. Further studies on the governance of data on common health conditions of older persons to inform clinical practice, health data retention schedules and destruction methods, and data governance policies and investments are recommended.

Acknowledgments

This scoping review was funded by the National Institutes of Health (NIH), University of the Philippines Manila through the NIH Commissioned Grants 2023 (project code: NIH 2023-027). The project team would like to thank and acknowledge the following: the NIH Executive Committee for funding this scoping review; Dr Shelley Ann de la Vega, director of the Institute on Aging, for her valuable support to the project especially during the funding application; Ms Vivian Biala for providing research grant administrative support; CAMP Research Committee for carefully reviewing and providing technical approval of the proposal; UP Manila Research Ethics Board for exempting the study from ethical review; experts consultants (Dr Alvin Marcelo, Dr Evelyn Hovenga, and Dr John Richard Hizon) for providing valuable insights based on their expertise; Mr Cristopher Naguit for helping the team process the salary and disbursement documents; and, finally, to the UP Manila channels for the processing of administrative,

legal, financial, and technical documents needed for this project. This project would not have been completed without the support of the above personalities and offices. All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided APF support for the publication of this paper.

Data Availability

All study data are present in this paper and the supplementary materials.

Authors' Contributions

AG performed protocol development, data collection, analysis, manuscript writing, review, and approval of the paper; AMB performed data collection, analysis, manuscript writing, review, and approval of the paper; and YP performed protocol development, data analysis, manuscript writing, review, and approval of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search terms used by database, summary of extracted entries from the included studies, and research gaps on data governance in the health care of older persons.
[DOCX File, 42 KB - aging_v8i1e73625_app1.docx]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[PDF File, 660 KB - aging_v8i1e73625_app2.pdf]

References

1. Weill P, Ross JW. IT Governance: How Top Performers Manage IT Decision Rights for Superior Results: Harvard Business School Press; 2004.
2. Data Administration Management Association. In: Mosley M, Brackett M, Earley S, Henderson D, editors. The DAMA Guide to the Data Management Body of Knowledge: DAMA-DMBOK Guide, 1st edition: Technics Publications, LLC; 2010.
3. Hovenga EJS, Grain H. Health data and data governance. Stud Health Technol Inform 2013;193(193):67-92. [Medline: 24018511]
4. Rosenbaum S. Data governance and stewardship: designing data stewardship entities and advancing data access. Health Serv Res 2010 Oct;45(5 Pt 2):1442-1455. [doi: 10.1111/j.1475-6773.2010.01140.x] [Medline: 21054365]
5. Micheli M, Ponti M, Craglia M, Berti Suman A. Emerging models of data governance in the age of datafication. Big Data Soc 2020 Jul;7(2). [doi: 10.1177/2053951720948087]
6. Hovenga E, Grain H. Quality data, design, implementation, and governance. In: Roadmap to Successful Digital Health Ecosystems: A Global Perspective: Elsevier; 2022:209-237. [doi: 10.1016/B978-0-12-823413-6.00013-6]
7. Oachs PK, Watters A, editors. Health Information Management: Concepts, Principles, and Practice, 6th edition: American Health Information Management Association; 2020.
8. Ngesimani NL, Ruhode E, Harpur PA. Data governance in healthcare information systems: a systematic literature review. S Afr J Inf Manag 2022 Aug 18;24(1). [doi: 10.4102/sajim.v24i1.1475]
9. Dickens A. From information to valuable asset: the commercialization of health data as a human rights issue. Health Hum Rights 2020 Dec;22(2):67-69. [Medline: 33390695]
10. Data Governance in Public Health: Pan American Health Organization; 2022. URL: https://iris.paho.org/handle/10665.2/56576 [accessed 2025-05-21]
11. Benfeldt Nielsen O. A comprehensive review of data governance literature. Selected Papers of the IRIS 2017 Dec 31(8):3 [FREE Full text]
12. Hovenga EJS. Impact of data governance on a nation's healthcare system building blocks. Stud Health Technol Inform 2013;193(193):24-66. [Medline: 24018510]
13. Healthcare data governance. American Health Information Management Association (AHIMA). 2022. URL: https://www.ahima.org/media/pmcb0fr5/healthcare-data-governance-practice-brief-final.pdf [accessed 2025-05-31]
14. Recommendation of the council on health data governance. OECD. 2025. URL: https://legalinstruments.oecd.org/public/doc/348/348.en.pdf [accessed 2025-06-12]

15. From the evolution of information systems for health to the digital transformation of the health sector. : Pan American Health Organization; 2021 URL: https://iris.paho.org/bitstream/handle/10665.2/53364/PAHOEIHIS210006_eng.pdf [accessed 2025-05-21]
16. Paparova D, Aanestad M, Vassilakopoulou P, Bahus MK. Data governance spaces: the case of a national digital service for personal health data. Inf Organ 2023 Mar;33(1):100451. [doi: 10.1016/j.infoandorg.2023.100451]
17. Zayas-Cabán T, Chaney KJ, Rucker DW. National health information technology priorities for research: a policy and development agenda. J Am Med Inform Assoc 2020 Apr 1;27(4):652-657. [doi: 10.1093/jamia/ocaa008]
18. Meeting on strengthening health information systems for sustainable development goals and UHC monitoring in the Western Pacific region. : World Health Organization Regional Office for the Western Pacific (WHO WPRO); 2019 URL: https://iris.who.int/handle/10665/325941?locale-attribute=fr&show=full [accessed 2025-05-21]
19. Williamson K, Nimegeer A, Lean M. Navigating data governance approvals to use routine health and social care data to evidence the hidden population with severe obesity: a case study from a clinical academic's perspective. J Res Nurs 2022 Nov;27(7):623-636. [doi: 10.1177/17449871221122040] [Medline: 36405806]
20. Good data: data sharing, public trust and willingness. World Economic Forum. 2021. URL: https://www.weforum.org/publications/good-data-sharing-data-and-fostering-public-trust-and-willingness [accessed 2025-05-21]
21. Why we need to leverage data for healthy ageing. World Economic Forum. 2021. URL: https://www.weforum.org/agenda/2021/10/we-need-to-leverage-data-for-healthy-ageing/ [accessed 2025-05-21]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
24. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
25. Nooteboom B. Learning by interaction: absorptive capacity, cognitive distance and governance. J Manage Governance 2000;4(1/2):69-92. [doi: 10.1023/A:1009941416749]
26. Liu Y, Zhang L, Yang Y, et al. A novel cloud-based framework for the elderly healthcare services using Digital Twin. IEEE Access 2019;7(7):49088-49101. [doi: 10.1109/ACCESS.2019.2909828]
27. Joseph RM, Knaggs RD, Coupland CAC, et al. Frequency and impact of medication reviews for people aged 65 years or above in UK primary care: an observational study using electronic health records. BMC Geriatr 2023 Jul 14;23(1):435. [doi: 10.1186/s12877-023-04143-2] [Medline: 37442984]
28. Pradhan A, Kuruganti U, Chester V. Biomechanical parameters and clinical assessment scores for identifying elderly fallers based on balance and dynamic tasks. IEEE Access 2020;8(8):193532-193543. [doi: 10.1109/ACCESS.2020.3033194]
29. Papoutsi C, Poots A, Clements J, Wyrko Z, Offord N, Reed JE. Improving patient safety for older people in acute admissions: implementation of the Frailsafe checklist in 12 hospitals across the UK. Age Ageing 2018 Mar 1;47(2):311-317. [doi: 10.1093/ageing/afx194] [Medline: 29315375]
30. Daniels H, Hollinghurst J, Fry R, et al. The value of routinely collected data in evaluating home assessment and modification interventions to prevent falls in older people: systematic literature review. JMIR Aging 2021 Apr 23;4(2):e24728. [doi: 10.2196/24728] [Medline: 33890864]
31. Fang L, Wu Y, Wu C, Yu Y. A nonintrusive elderly home monitoring system. IEEE Internet Things J 2021;8(4):2603-2614. [doi: 10.1109/JIOT.2020.3019270] [Medline: 37974935]
32. Bernsmed K. Accountable health care service provisioning in the cloud. 2014 Presented at: 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC); Dec 8-11, 2014; London, England, United Kingdom p. 902-907. [doi: 10.1109/UCC.2014.147]
33. Thorogood A, Mäki-Petäjä-Leinonen A, Brodaty H, Global Alliance for Genomics and Health, Ageing and Dementia Task Team, et al. Consent recommendations for research and international data sharing involving persons with dementia. Alzheimers Dement 2018 Oct;14(10):1334-1343. [doi: 10.1016/j.jalz.2018.05.011] [Medline: 30293575]
34. Kiefer R, Abid M, Steen J, Ardali MR, Amjadian E. A catalog of public glaucoma datasets for machine learning applications: a detailed description and analysis of public glaucoma datasets available to machine learning engineers tackling glaucoma-related problems using retinal fundus images and OCT images. : ACM; 2023 Presented at: 2023 the 7th International Conference on Information System and Data Mining (ICISDM); May 10-12, 2023; Atlanta, United States p. 24-31. [doi: 10.1145/3603765.3603779]
35. Billis AS, Papageorgiou EI, Frantzidis CA, Tsatali MS, Tsolaki AC, Bamidis PD. A decision-support framework for promoting independent living and ageing well. IEEE J Biomed Health Inform 2015 Jan;19(1):199-209. [doi: 10.1109/JBHI.2014.2336757] [Medline: 25073180]
36. Chung-Chih L, Ping-Yeh L, Po-Kuan L, Guan-Yu H, Wei-Lun L, Ren-Guey L. A healthcare integration system for disease assessment and safety monitoring of dementia patients. IEEE Trans Inform Technol Biomed 2008;12(5):579-586. [doi: 10.1109/TITB.2008.917914]

37. Wieland-Jorna Y, Verheij RA, Francke AL, et al. Setting up a governance framework for secondary use of routine health data in nursing homes: development study using qualitative interviews. J Med Internet Res 2023 Jan 25;25(25):e38929. [doi: 10.2196/38929] [Medline: 36696162]
38. Dahlberg T. A framework for and issues in the management and governance of data on elderly citizens (ongoing res.). Presented at: 8th International Conference on Theory and Practice of Electronic Governance; Oct 27-30, 2014; Guimaraes, Portugal. [doi: 10.1145/2691195.2691295]
39. Zhou S, Ni Z, Ogihara A, Wang X. Behavioral patterns of supply and demand sides of health services for the elderly in sustainable digital transformation: a mixed methods study. Int J Environ Res Public Health 2022 Jul 5;19(13):8221. [doi: 10.3390/ijerph19138221] [Medline: 35805878]
40. Zhang T, Sodhro AH, Luo Z, et al. A joint deep learning and internet of medical things driven framework for elderly patients. IEEE Access 2020;8(8):75822-75832. [doi: 10.1109/ACCESS.2020.2989143]
41. Lin D, Yao C, Min W, et al. A novel fall detection framework with age estimation based on cloud-fog computing architecture. IEEE Sensors J 2024 Feb;24(3):3058-3071. [doi: 10.1109/JSEN.2023.3334555]
42. Khalid S. Towards data governance for international dementia care mapping (DCM). a study proposing DCM data management through a data warehousing approach [Master's thesis]. : The University of Bradford; 2011 Nov 18 URL: https://bradscholars.brad.ac.uk/handle/10454/5226 [accessed 2025-05-21]
43. Lin CH, Wang FC, Kuo TY, Huang PW, Chen SF, Fu LC. Early detection of Parkinson's disease by neural network models. IEEE Access 2022;10(10):19033-19044. [doi: 10.1109/ACCESS.2022.3150774]
44. Buleje I, Siu VS, Hsieh KY, et al. A versatile data fabric for advanced IoT-based remote health monitoring. Presented at: 2023 IEEE International Conference on Digital Health (ICDH); Jul 2-8, 2023; Chicago, IL, United States. [doi: 10.1109/ICDH60066.2023.00021]
45. Béland F, Bergman H, Lebel P, et al. A system of integrated care for older persons with disabilities in Canada: results from a randomized controlled trial. J Gerontol 2006 Apr 1;61(4):367-373. [doi: 10.1093/gerona/61.4.367]
46. Greene BR, McManus K, Redmond SJ, Caulfield B, Quinn CC. Digital assessment of falls risk, frailty, and mobility impairment using wearable sensors. NPJ Digit Med 2019;2(2):125. [doi: 10.1038/s41746-019-0204-z] [Medline: 31840096]
47. Shulman MA, Wallace S, Gilbert A, Reilly JR, Kasza J, Myles PS. Predicting death or disability after surgery in the older adult. Anesthesiology 2023;139(4):420-431. [doi: 10.1097/ALN.0000000000004683]
48. Kim JY, Liu N, Tan HX, Chu CH. Unobtrusive monitoring to detect depression for elderly with chronic illnesses. IEEE Sensors J 2017 Sep 1;17(17):5694-5704. [doi: 10.1109/JSEN.2017.2729594]
49. Choudhury A, Renjilian E, Asan O. Use of machine learning in geriatric clinical care for chronic diseases: a systematic literature review. JAMIA Open 2020 Oct;3(3):459-471. [doi: 10.1093/jamiaopen/ooaa034] [Medline: 33215079]
50. Elbattah M, Molloy O. Coupling simulation with machine learning: a hybrid approach for elderly discharge planning. : ACM; 2016 Presented at: Proceedings of the 2016 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation; May 15-18, 2016; Banff, Alberta, Canada p. 47-56. [doi: 10.1145/2901378.2901381]
51. Hu X, Chu D, Li Z, et al. Coarse-to-fine activity annotation and recognition algorithm for solitary older adults. IEEE Access 2020;8(8):4051-4064. [doi: 10.1109/ACCESS.2019.2962843]
52. Subbe CP, Goulden N, Mawdsley K, Smith R. Anticipating care needs of patients after discharge from hospital: frail and elderly patients without physiological abnormality on day of admission are more likely to require social services input. Eur J Intern Med 2017 Nov;45(45):74-77. [doi: 10.1016/j.ejim.2017.09.029] [Medline: 28974330]
53. Unleashing the power of big data for Alzheimer's disease and dementia research: main points of the OECD expert consultation on unlocking global collaboration to accelerate innovation for Alzheimer's disease and dementia. : OECD Publishing; 2014 Mar URL: http://dx.doi.org/10.1787/5jz73kvmvbwb-en [accessed 2025-06-06]
54. Lin KC, Wai RJ. A feasible fall evaluation system via artificial intelligence gesture detection of gait and balance for sub-healthy community- dwelling older adults in Taiwan. IEEE Access 2021;9(9):146404-146413. [doi: 10.1109/ACCESS.2021.3123297]
55. Taramasco C, Rodenas T, Martinez F, et al. A novel monitoring system for fall detection in older people. IEEE Access 2018;6(6):43563-43574. [doi: 10.1109/ACCESS.2018.2861331]
56. Schutz N, Botros A, Hassen SB, et al. A sensor-driven visit detection system in older adults' homes: towards digital late-life depression marker extraction. IEEE J Biomed Health Inform 2022 Apr;26(4):1560-1569. [doi: 10.1109/JBHI.2021.3114595] [Medline: 34550895]
57. Austin J, Dodge HH, Riley T, Jacobs PG, Thielke S, Kaye J. A smart-home system to unobtrusively and continuously assess loneliness in older adults. IEEE J Transl Eng Health Med 2016;4(4):2800311. [doi: 10.1109/JTEHM.2016.2579638] [Medline: 27574577]
58. Rajamäki J, Gioulekas F, Rocha PAL, Garcia XDT, Ofem P, Tyni J. ALTAI tool for assessing ai-based technologies: lessons learned and recommendations from SHAPES pilots. Healthcare (Basel) 2023 May 17;11(10):1454. [doi: 10.3390/healthcare11101454]
59. Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Büla CJ, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. IEEE Trans Biomed Eng 2003 Jun;50(6):711-723. [doi: 10.1109/TBME.2003.812189] [Medline: 12814238]

60. Elhussein L, Jödicke AM, He Y, et al. Characterising complex health needs and the use of preventive therapies in the older population: a population-based cohort analysis of UK primary care and hospital linked data. BMC Geriatr 2023 Jan 31;23(1):58. [doi: 10.1186/s12877-023-03770-z] [Medline: 36721104]
61. Manh Do H, Sheng W, Harrington EE, Bishop AJ. Clinical screening interview using a social robot for geriatric care. IEEE Trans Automat Sci Eng 2021;18(3):1229-1242. [doi: 10.1109/TASE.2020.2999203]
62. Spinsante S, Poli A, Mongay Batalla J, et al. Clinically-validated technologies for assisted living. J Ambient Intell Human Comput 2023 Mar;14(3):2095-2116. [doi: 10.1007/s12652-021-03419-y]
63. Bajenaru L, Marinescu IA, Dobre C, Tomescu M, Herghelegiu AM. Clinically-validated technologies for older adults' quality of life self-management: vINCI ecosystem. Presented at: 2022 14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI); Jun 30 to Jul 1, 2022; Ploiesti, Romania. [doi: 10.1109/ECAI54874.2022.9847465]
64. Delmastro F, Martino FD, Dolciotti C. Cognitive training and stress detection in MCI frail older people through wearable sensors and machine learning. IEEE Access 2020;8(8):65573-65590. [doi: 10.1109/ACCESS.2020.2985301]
65. Yao L, Sheng QZ, Li X, et al. Compressive representation for device-free activity recognition with passive RFID signal strength. IEEE Trans Mobile Comput 2018 Feb;17(2):293-306. [doi: 10.1109/TMC.2017.2706282]
66. Moreno PA, Hernando ME, Gómez EJ. Design and technical evaluation of an enhanced location-awareness service enabler for spatial disorientation management of elderly with mild cognitive impairment. IEEE J Biomed Health Inform 2015 Jan;19(1):37-43. [doi: 10.1109/JBHI.2014.2327638] [Medline: 25486651]
67. Andrew NE, Beare R, Ravipati T, et al. Developing a linked electronic health record derived data platform to support research into healthy ageing. Int J Popul Data Sci 2023;8(1):2129. [doi: 10.23889/ijpds.v8i1.2129] [Medline: 37670961]
68. Lin YH, Mao HF, Lin KN, Tang YL, Yang CL, Chou JJ. Development and evaluation of a computer game combining physical and cognitive activities for the elderly. IEEE Access 2020;8(8):216822-216834. [doi: 10.1109/ACCESS.2020.3041017]
69. McManus K, Greene BR, Ader LGM, Caulfield B. Development of data-driven metrics for balance impairment and fall risk assessment in older adults. IEEE Trans Biomed Eng 2022 Jul;69(7):2324-2332. [doi: 10.1109/TBME.2022.3142617] [Medline: 35025734]
70. Wu J, Chang L, Yu G. Effective data decision-making and transmission system based on mobile health for chronic disease management in the elderly. IEEE Syst J 2021;15(4):5537-5548. [doi: 10.1109/JSYST.2020.3024816]
71. Chen T, Ding Z, Li B. Elderly fall detection based on improved YOLOv5s network. IEEE Access 2022;10(10):91273-91282. [doi: 10.1109/ACCESS.2022.3202293]
72. Doyle J, Caprani N, Kealy A, Bond R, Komaba Y, Inomata A. Healthcare professionals views on technology to support older adults transitioning from hospital to home. Presented at: Proceedings of the 30th International BCS Human Computer Interaction Conference; Jul 11-15, 2016; Poole, United Kingdom. [doi: 10.14236/ewic/HCI2016.34]
73. Ayena JC, Zaibi H, Otis MJD, Menelas BAJ. Home-based risk of falling assessment test using a closed-loop balance model. IEEE Trans Neural Syst Rehabil Eng 2016 Dec;24(12):1351-1362. [doi: 10.1109/TNSRE.2015.2508960] [Medline: 26685258]
74. Muheidat F, Tawalbeh LA. In-home floor based sensor system-smart carpet- to facilitate healthy aging in place (AIP). IEEE Access 2020;8(8):178627-178638. [doi: 10.1109/ACCESS.2020.3027535]
75. Kim J, Cheon S, Lim J. IoT-based unobtrusive physical activity monitoring system for predicting dementia. IEEE Access 2022;10(10):26078-26089. [doi: 10.1109/ACCESS.2022.3156607]
76. Yu L, Chan WM, Zhao Y, Tsui KL. Personalized health monitoring system of elderly wellness at the community level in Hong Kong. IEEE Access 2018;6(6):35558-35567. [doi: 10.1109/ACCESS.2018.2848936]
77. van der Meer HG, Wouters H, Teichert M, et al. Feasibility, acceptability and potential effectiveness of an information technology-based, pharmacist-led intervention to prevent an increase in anticholinergic and sedative load among older community-dwelling individuals. Ther Adv Drug Saf 2019;10(10):2042098618805881. [doi: 10.1177/2042098618805881] [Medline: 31019675]
78. Nath RK, Thapliyal H. Smart wristband-based stress detection framework for older adults with cortisol as stress biomarker. IEEE Trans Consumer Electron 2021;67(1):30-39. [doi: 10.1109/TCE.2021.3057806]
79. Lutomski JE, Baars MAE, Schalk BWM, et al. The development of the Older Persons and Informal Caregivers Survey Minimum DataSet (TOPICS-MDS): a large-scale data sharing initiative. PLoS One 2013;8(12):e81673. [doi: 10.1371/journal.pone.0081673] [Medline: 24324716]
80. Lewis ET, Williamson M, Lewis LP, et al. The feasibility of deriving the electronic frailty index from Australian general practice records. Clin Interv Aging 2022;17(17):1589-1598. [doi: 10.2147/CIA.S384691] [Medline: 36353269]
81. Zherebtsov EA, Zharkikh EV, Loktionova YI, et al. Wireless dynamic light scattering sensors detect microvascular changes associated with ageing and diabetes. IEEE Trans Biomed Eng 2023;70(11):3073-3081. [doi: 10.1109/TBME.2023.3275654]
82. Lin CC, Chiu MJ, Hsiao CC, Lee RG, Tsai YS. Wireless health care service system for elderly with dementia. IEEE Trans Inform Technol Biomed 2006;10(4):696-704. [doi: 10.1109/TITB.2006.874196]

83. Garcia AP, De La Vega SF, Mercado SP. Health information systems for older persons in select government tertiary hospitals and health centers in the Philippines: cross-sectional study. J Med Internet Res 2022 Feb 14;24(2):e29541. [doi: 10.2196/29541] [Medline: 35156927]
84. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies: World Health Organization; 2010. URL: https://iris.who.int/handle/10665/258734 [accessed 2025-05-21]
85. Cave A. Exploring strategies for implementing data governance practices [dissertation]. : Walden University; 2017 Jan 1 URL: https://scholarworks.waldenu.edu/dissertations/4206
86. Health Data Governance. 2025. URL: https://healthdatagovernance.org/ [accessed 2025-05-21]
87. Aïdoud A, Gana W, Poitau F, et al. High prevalence of geriatric conditions among older adults with cardiovascular disease. J Am Heart Assoc 2023 Jan 17;12(2):e026850. [doi: 10.1161/JAHA.122.026850] [Medline: 36628962]
88. Gómez-Ramos JJ, González-Guerra M, Dávalos-Rodríguez IP, Pérez-Ruíz ME, Peña-Durán E, Marín-Medina A. Geriatric syndromes and their relationship with mortality in a population of Mexican older adults aged 65 and over, admitted to the emergency department of a second-level care hospital. Healthcare (Basel) 2024 Jun 8;12(12):1166. [doi: 10.3390/healthcare12121166]
89. Wang LY, Hu ZY, Chen HX, Tang ML, Hu XY. Multiple geriatric syndromes in community-dwelling older adults in China. Sci Rep 2024 Feb 12;14(1):3504. [doi: 10.1038/s41598-024-54254-y] [Medline: 38347070]
90. Ward KT, Reuben DB, Schmader KE, Givens J. Comprehensive geriatric assessment. UpToDate. 2024 May 17. URL: https://www.uptodate.com/contents/comprehensive-geriatric-assessment [accessed 2025-06-12]
91. Earley S, Data Management Association. The DAMA Dictionary of Data Management: Technics Publications; 2011.
92. Cleven A, Wortmann F. Uncovering four strategies to approach master data management. Presented at: 2010 43rd Hawaii International Conference on System Sciences; Jan 5-8, 2010; Honolulu, HI, United States. [doi: 10.1109/HICSS.2010.488]
93. El-Jardali F, Bou-Karroum L, Jabbour M, et al. Digital health in fragile states in the Middle East and North Africa (MENA) region: a scoping review of the literature. PLoS One 2023;18(4):e0285226. [doi: 10.1371/journal.pone.0285226]
94. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
95. Monlezun DJ, Omutoko L, Oduor P, et al. Digitalization of health care in low- and middle-income countries. Bull World Health Organ 2025 Feb 1;103(2):148-154. [doi: 10.2471/BLT.24.291643] [Medline: 39882493]
96. A. Alampay E. Open data governance in the time of the pandemic. : ACM Presented at: 14th International Conference on Theory and Practice of Electronic Governance; Oct 6-8, 2021; Athens, Greece p. 455-462. [doi: 10.1145/3494193.3494255]
97. Koltay T. Data governance, data literacy and the management of data quality. IFLA J 2016 Dec;42(4):303-312. [doi: 10.1177/0340035216672238]
98. Delacroix S, Lawrence ND. Bottom-up data trusts: disturbing the ‘one size fits all’ approach to data governance. Int Data Privacy Law 2019 Oct 1:ipz014. [doi: 10.1093/idpl/ipz014]
99. Yukhno A. Digital transformation: exploring big data governance in public administration. Public Organ Rev 2024 Mar;24(1):335-349. [doi: 10.1007/s11115-022-00694-x]
100. Platt JE, Raj M, Wienroth M. An analysis of the learning health system in its first decade in practice: scoping review. J Med Internet Res 2020 Mar 19;22(3):e17026. [doi: 10.2196/17026] [Medline: 32191214]
101. Zurynski Y, Smith CL, Vedovi A, et al. Mapping the learning health system: a scoping review of current evidence. In: NHRMC Partnership Centre for Health System Sustainability: Australian Institute of Health Innovation; 2020.

Abbreviations

ACM: Association for Computing Machinery
PAHO: Pan American Health Organization
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by PH Chau; submitted 07.03.25; peer-reviewed by MA Zoubi, S Mohanadas, T Adebambo; revised version received 26.04.25; accepted 30.04.25; published 27.06.25.

Please cite as:
Garcia A, Balingbing AM, Palad Y
Exploring Literature on Data Governance in the Health Care of Older Persons: Scoping Review
JMIR Aging 2025;8:e73625
URL: https://aging.jmir.org/2025/1/e73625
doi:10.2196/73625

Review

Nonpharmacological Multimodal Interventions for Cognitive Functions in Older Adults With Mild Cognitive Impairment: Scoping Review

Raffy Chi-Fung Chan[1*], MA; Joson Hao-Shen Zhou[1*], MMedSc; Yuan Cao[2], PhD; Kenneth Lo[3,4], PhD; Peter Hiu-Fung Ng[1], PhD; David Ho-Keung Shum[1,4], PhD; Arnold Yu-Lok Wong[1,4], PhD

[1]Department of Rehabilitation Sciences, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Department of Social Work and Social Administration, University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Department of Food Science and Nutrition, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[4]Research Institute for Smart Ageing, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Arnold Yu-Lok Wong, PhD
Department of Rehabilitation Sciences
Hong Kong Polytechnic University
QT522, 5/F, Core T
The Hong Kong Polytechnic University
Hong Kong
China (Hong Kong)
Phone: 852 2766 6741
Email: arnold.wong@polyu.edu.hk

Abstract

Background: As the global population ages, the prevalence of dementia is expected to rise significantly. To alleviate the burden on health care systems and the economy, it is essential to develop effective strategies to enhance cognitive function in older adults. Previous studies have shown that combined nonpharmacological interventions can improve cognition across various domains in older individuals. However, there is no established gold standard for the exact combination and duration of these interventions, which makes it challenging to assess their overall effectiveness.

Objective: Given the diversity of nonpharmacological multimodal interventions aimed at preventing cognitive decline in older adults with mild cognitive impairment (MCI), this scoping review sought to identify and summarize the characteristics and outcomes of these interventions.

Methods: We adhered to the Arksey and O'Malley methodological framework and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) and searched 4 electronic databases (MEDLINE, PsycINFO, CINAHL, and Web of Science) systematically on July 6, 2023, and updated the search on April 17, 2024, using specific terms and keywords.

Results: This review included 45 studies from 18 countries with 4705 participants from 2014 to 2024 encompassing different combinations of physical training (PT), cognitive training (CT), nutrition intervention, psychosocial intervention, social activities, and electrical stimulation. There is a growing numbers of studies combining PT and CT for MCI treatment, with additional modalities often added to address various aspects of the condition. Compared to single-modal interventions and usual care, multimodal approaches demonstrated significantly better improvements in cognition domains such as attention, global cognition, executive function, memory, processing speed, and verbal fluency. Technology has been instrumental in delivering these interventions and enhancing the effects of PT and CT. Multimodal interventions also show promise in terms of acceptability and user experience, which can improve treatment adherence.

Conclusions: Research is limited regarding the cost-effectiveness and optimal dosage of these interventions, making it difficult to assess the additional benefits of incorporating more modalities. Future research should examine the long-term effects of incorporating multiple modalities, using standardized MCI criteria and outcome measures.

(JMIR Aging 2025;8:e70291) doi:10.2196/70291

KEYWORDS

mild cognitive impairment; multimodal intervention; prevention; randomized controlled trial; cognitive decline

Introduction

Overview

Mild cognitive impairment (MCI) represents a transitional stage between normal aging and dementia, characterized by cognitive decline greater than expected for age, impacting one or more domains such as attention, memory, orientation, executive functioning, language, and visuospatial skills [1,2]. However, the decline does not significantly interfere with daily activities [3]. The lifetime prevalence of MCI in individuals aged above 60 years is estimated to be 15% to 20% [4], with this prevalence increasing with age. The annual transition rate from MCI to various subtypes of dementia ranges from 10% to 15% and can reach up to 25.2% for adults aged 80 to 84 years [5,6]. Older adults with MCI are 46% more likely to develop dementia within 3 years compared to just 3% in the normal-aging population [7].

The trajectory for individuals with MCI may vary, leading to dementia, stable cognition, or a return to normal cognition function [6]. Untreated MCI may progress to dementia, a neurodegenerative disease that significantly impacts daily functions and affects one's physical, psychological, social, and economic aspects. Dementia also directly affects one's caregivers, families, and society [8]. With the global population of individuals aged 60 years and above projected to double from 1 billion in 2020 to 2 billion by 2050 [9], there is mounting concern regarding the rising prevalence of dementia and the urgent need for preventive measures to address the associated social and economic burdens. Factors, such as modifiable risk factors, genetics, and interventions, can affect MCI progression [10]. Therefore, understanding various interventions and their effectiveness in preventing progression from MCI to dementia is crucial given the rising prevalence of dementia among older adults.

Currently, there is no gold standard for treatments or interventions to manage MCI. The American Academy of Neurology guidelines indicate insufficient empirical evidence to support pharmacological treatments for MCI in older adults [11]. Conversely, numerous studies have advocated nonpharmacological interventions, such as physical training (PT) or cognitive training (CT), as effective strategies for managing MCI [11-13].

Nonpharmacological Interventions

PT, including aerobic, strengthening, and balance exercises, has been shown to stimulate norepinephrine release in the brain, promote brain plasticity, increase brain volume, and enhance cerebral blood flow [14]. These effects are crucial for improving cognition, mood, and physiological abilities [15]. Combining aerobic and strengthening exercises has been identified as particularly effective for cognitive improvement [16].

On the contrary, CT primarily uses cognitive stimulation and repetitive tasks to enhance various cognitive domains, particularly memory, attention, and executive function [17]. The effectiveness of CT in improving cognition lies in its capacity to strengthen the functioning and plasticity of neural networks and cognitive reserves [17]. Research suggested that memory-focused CT increased activation and connectivity in the frontal, temporal, and occipital brain regions [18]. These areas are crucial for memory, motor function, processing, attention, language, mood, and problem-solving. Therefore, CT shows promise for improving overall cognition in older adults with MCI.

Given that both PT and CT can enhance brain plasticity and stimulate brain regions responsible for various cognitive functions, their combined application, whether delivered separately or through dual tasking, is the most common approach for managing cognitive impairment in MCI. This multifaceted approach effectively targets different aspects of cognitive decline, countering cognitive decay and neurodegeneration.

Growing evidence suggests that engaging in more social activities (SA) can lower the risk of cognitive decline in individuals with MCI [5], making SA a potential key modality for MCI management. SA involves participation in activities that allow interactions or engagements with others [5]. Karp et al [19] found that older adults with MCI who participated in a broader range of activities, including mental stimulation, physical activities, and social recreation, had a lower risk of developing dementia than those who participated in fewer or no such activities. On the other hand, beyond PT, CT, and SA, emerging evidence suggests that modifiable lifestyle factors, including diet and nutritional intervention (NI), electrical stimulation, or psychosocial intervention (PI) may also improve cognitive functions in this population [10,20].

Multimodal Interventions

A multimodal intervention, integrating various methods such as PT, CT, SA, NI, electrical stimulation, and PI, either sequentially or simultaneously, addresses cognitive decline across different domains of MCI. This approach has been proven effective in managing MCI in older adults, enhancing cognitive abilities, mood, sleep, activities of daily living, functional capacities, and physical abilities, with benefits lasting up to 2 years [14]. A systematic review and meta-analysis [21] showed that combining 2 or more interventions had small to medium effects on global cognition, memory, executive function, and verbal fluency, demonstrating a synergistic effect. Studies also showed that multimodal interventions outperformed single-modal interventions in managing MCI [15,16]. Another systematic review and meta-analysis [22] also supported that combined PT and CT had a small to medium effect on global cognition than various types of cognitive-only interventions in older adults with MCI. However, these reviews solely focused on the effects of combined PT and CT [22] or compared the

effectiveness of multimodal interventions only with single-modal interventions [21], thus lacking direct comparisons between various types of multimodal interventions to inform researchers or clinicians regarding which combinations of multimodal interventions would yield better results. In addition, previous reviews did not consider the user experience of multimodal interventions, which limits the clinical applicability of their findings and important factors to determine the feasibility of multimodal interventions. Thus, a comprehensive evaluation of multimodal interventions should include both feasibility and user experience to optimize benefits.

Optimizing Multimodal Interventions: Technology, Dosage, and Cost-Effectiveness

The increasing adoption of digital health care technology in managing MCI [23] could enhance the delivery and reduce the costs of interventions, especially benefiting those in remote areas [24]. Given the high prevalence and economic burden of MCI on communities and health care systems, technology such as the use of computerized CT (CCT) can help mitigate costs associated with nonpharmacological interventions, addressing the shortage of trained professionals. While previous systematic reviews and meta-analyses have acknowledged the role of technology in MCI management, they mainly focused on CCT [24-26] without considering the broader potential of technology beyond CCT.

Addressing the impact of different dosages of nonpharmacological interventions for MCI is crucial for developing effective and sustainable therapeutic strategies. This involves determining the optimal frequency, intensity, type, and duration of interventions to maximize cognitive benefits while balancing time and financial costs [27]. A study [27] highlighted that CT, PT, NI, and the majority of combined PT and CT significantly improved cognition in individuals with MCI, with particularly effective doses being 1 to 2 sessions per week with 60 to 120 minutes per session and interventions lasting over 12 weeks. However, as this study focused on only one type of multimodal intervention (combined PT and CT), the dosage effects of other nonpharmacological multimodal interventions remain uncertain. Therefore, this scoping review aimed to summarize dosage effects from the literature on multimodal interventions.

In addition, managing cognitive decline imposes a significant global economic burden. The cost-effectiveness of multimodal interventions varies by region because of the differences in health care resources. Although previous research supported the effectiveness of multimodal interventions for cognitive decline, their adoption of multiple modalities often increases cost. However, no prior review has summarized the cost-effectiveness of these interventions, so this scoping review aimed to address this important gap to inform clinical decision-making.

Rationale for a Scoping Review

Considering the variety of multimodal MCI interventions, both with and without technology, and the diverse methodologies and research focus of the existing literature, a scoping review is warranted to identify the current research gaps to inform future research and clinical practice. The complexity and heterogeneity of the interventions, coupled with the rapidly evolving nature of this field, make a systematic review less feasible for comprehensively mapping the use of different types of multimodal interventions and investigating the current research trends. As such, this scoping review aimed to map and describe the latest development in MCI interventions and provide researchers and clinicians with insights into current trends and limitations of the existing approaches and studies. Specifically, this review mapped the current landscape of nonpharmacological multimodal interventions for older adults with MCI, identified and summarized the components of these interventions, research trends, and the use of different outcome measures. It aimed to enhance the understanding of MCI management and provide future research directions on multimodal interventions.

Methods

Overview

This review adhered to the Arksey and O'Malley methodological framework [28] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [29]. The protocol was registered on the Open Science Framework platform [30]. The review process involved five key steps: (1) formulating the research questions; (2) devising the search strategy; (3) identifying and selecting relevant studies; (4) data charting; and (5) synthesizing and presenting findings.

Identified Research Questions

This review explored the following research questions:

1. What are the research trends in multimodal interventions for older adults with MCI?
2. What components were included in these multimodal interventions? What results have been reported?
3. How cost-effective were the identified multimodal interventions?
4. What role did technology play in these interventions for older adults with MCI?
5. What insights were available regarding the acceptability, user experiences, and dose responses of these interventions?

Identifying and Selecting Relevant Studies

We conducted a search across 4 databases—MEDLINE, PsycINFO, CINAHL, and Web of Science—on July 6, 2023, and updated the search on April 17, 2024, using specific Medical Subject Headings terms and keywords such as "combine," "multi," "dual," "mix interventions, and "mild cognitive impairment" (Multimedia Appendix 1). Although quality assessment is optional for scoping reviews [31], this review included only randomized controlled trials to enhance study quality. Two independent reviewers (JHSZ and RCFC) screened titles, abstracts, and full texts against the eligibility criteria (Textbox 1). Disagreements were resolved by consensus or by consulting a third reviewer (AYLW). The interrater reliability of the screening process as measured by the kappa coefficient was 0.87.

Textbox 1. Selection criteria for the scoping review.

Inclusion criteria

- Participants: diagnosed with mild cognitive impairment (MCI) by clinicians or psychologists using well-established criteria
- Intervention: at least one combination of a nonpharmacological multimodal intervention in managing older adults with MCI
- Control: received at least one or multiple forms of an intervention, a placebo or sham training, health education, or treatment as usual
- Outcome: must use at least one well-established measurement for testing cognitive outcomes
- Study design: must be an experimental study (randomized controlled trial or quasi-experimental study)
- Other: full-text and peer-reviewed study written in English

Exclusion criteria

- Participants: diagnosed with dementia or the cognitive impairment resulted from drug use or psychiatric or other neurological disorders (eg, bipolar disorder, schizophrenia, stroke, Parkinson disease, or epilepsy)
- Intervention: pharmacological interventions or nonpharmacological experimental studies with a single-modal intervention
- Study design: systematic reviews, scoping reviews, opinion letters, conference proceedings, dissertations, and research design protocols
- Other: gray literature

Definition of Types of Intervention and Control Groups

The included studies featured diverse intervention components and control groups. Detailed operational definitions for these interventions and control groups are provided in Textbox 2.

Textbox 2. Operational definition of the intervention and control groups in the included studies.

Interventions

- Physical training: activities, exercises, or training that required older adults to do physical activities with or without the guidance or supervision of a professional trainer or clinician
- Cognitive training: activities or training that used a standardized systematic cognitive stimulation, rehabilitation task, and training to improve cognitive function
- Nutrition intervention: the use of any type of dietary supplement, including herb extract, or any form of dietary counseling
- Psychosocial intervention: the use of activities, training, counseling, therapy, or education that aimed to improve psychological well-being, including music therapy and mindfulness
- Social activities: activities that encourage social engagement or facilitate social interaction between older adults
- Electric current stimulation: the use of current stimulation including transcranial alternating current stimulation and transcranial direct current stimulation

Control groups

- Active control: a group of participants who received at least one form of intervention
- Placebo control: a group of participants who received a placebo or sham training
- Health education control: a group of participants who received health education
- Inactive control: a group of participants who received no additional treatment, treatment as usual, or only health advice

Data Extraction

Relevant data, including authors, publication year, country, place of recruitment, diagnostic criteria, participants, intervention types, outcomes, treatment frequency and duration, follow-up time points, use of measurement tools, control group characteristics, results, and interpretations of findings, were extracted by 2 independent reviewers (JHSZ and RCFC). All extracted findings were compared. Discrepancies were reconciled by consensus or through consultation with a third reviewer (AYLW).

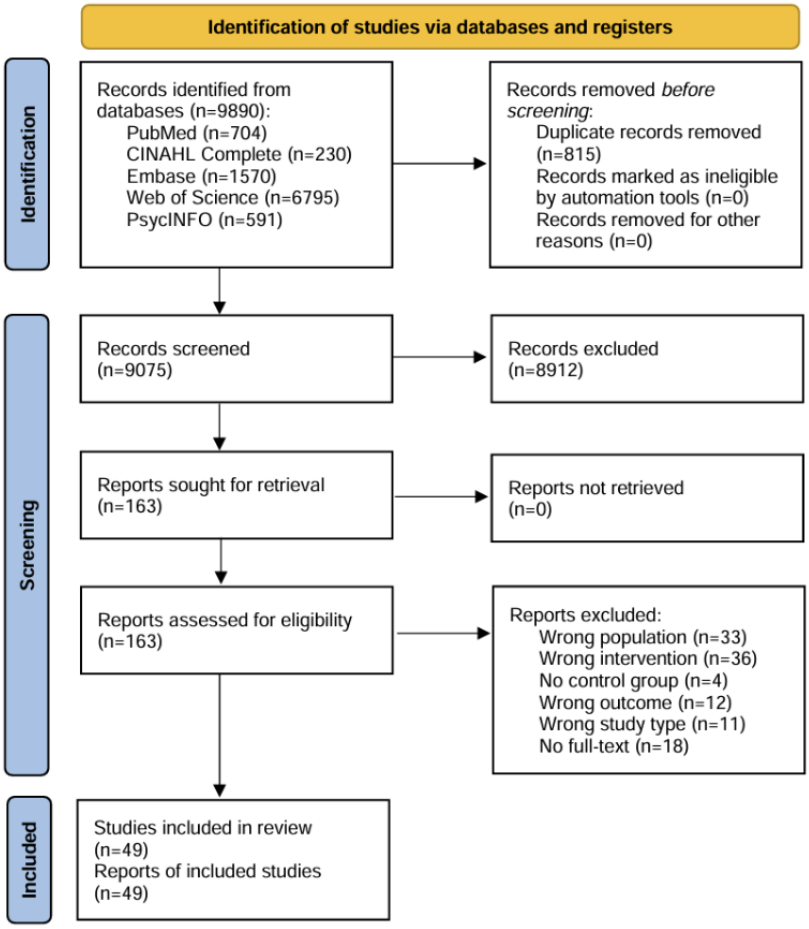
Results

Study Selection and Characteristics

The initial search identified 9890 articles. After removing 815 duplicates, 9075 titles and abstracts were screened; 163 out of 9075 titles and abstracts were selected for full-text screening. Studies based on the same cohort were counted as a single study, including 4 studies by Hagovská and Nagyova [32], Hagovská and Olekszyova [33,34], and Hagovská et al [35] and 2 studies by Liao et al [36,37]. Exclusions were made for reasons

including incorrect target population (n=33), absence of multimodal intervention (n=36), lack of a control group (n=4), different study outcomes (n=12), absence of cognitive outcomes (n=11), and unavailability of full text (n=18). Ultimately, 45 studies from 49 articles published between 2014 and 2024, encompassing a total of 4705 participants, were included in this review. A detailed description of the study selection is presented in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of literature search and screening.



The characteristics of the included studies, including participant characteristics, intervention tasks, treatment duration and frequency, and assessment time points, are presented in Multimedia Appendix 2. These studies used various criteria to diagnose MCI, such as the Peterson criteria (n=14) [15,16,38-49], the Albert criteria (n=5) [50-54], and cutoff scores from established screening tools like the Alzheimer Disease Assessment Scale–Cognitive subscale (cutoff score not provided) [38,40], Mini-Mental State Examination between 20 and 27 [55], Montreal Cognitive Assessment with scores up to 28 [15,23,36,37,52,56-65], Clinical Dementia Rating scores from 0.5 to less than 1 [45,60,66], or direct diagnoses from professional psychologists. Table 1 provides a detailed description of the diagnostic criteria used for a comprehensive reference.

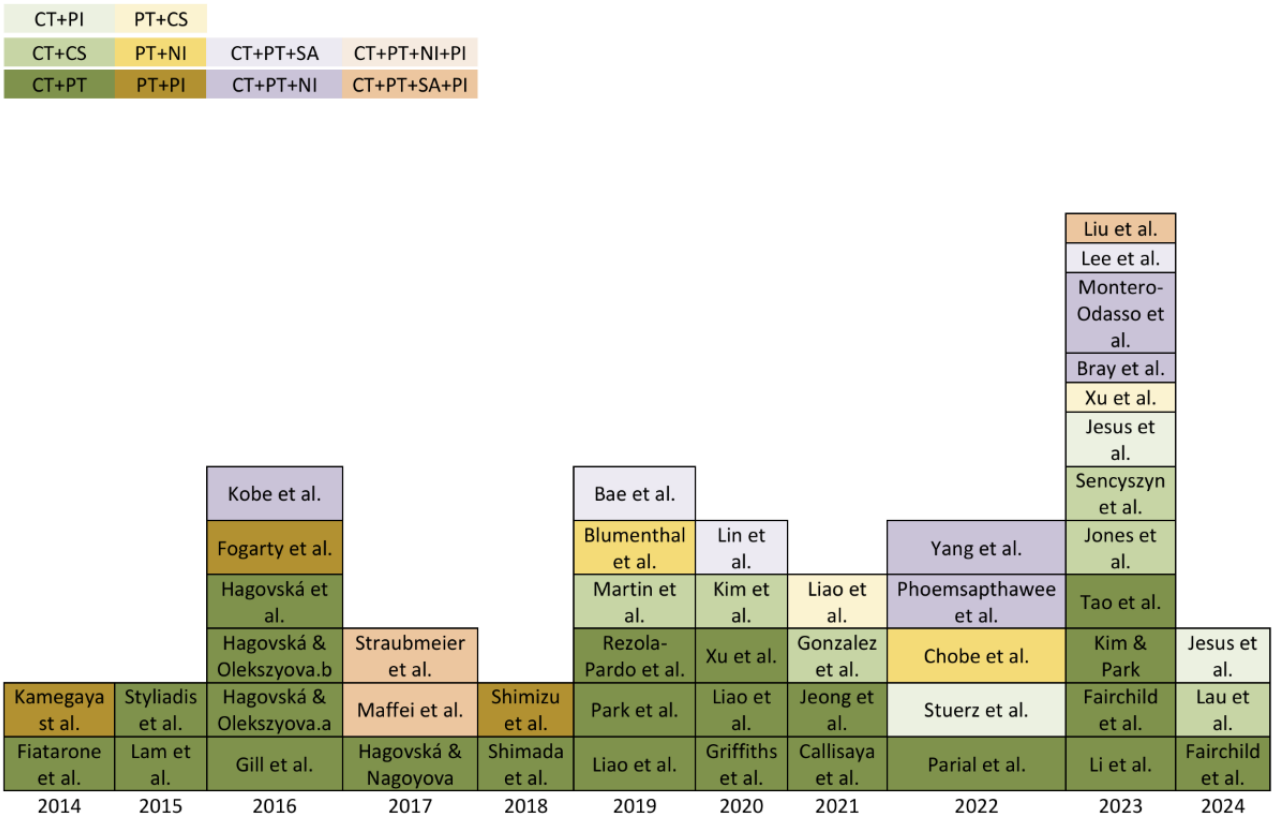
Table 1. Diagnostic criteria used in the included studies.

Criteria	Definition
Peterson criteria [11]	• Self-report cognitive decline • Objective cognitive impairment compared with age • No impact on daily functioning • No dementia
Albert criteria [3]	• Cognitive change reported by the patient, informant, or clinician • Objective evidence of cognitive impairment in one or more domains (typically 1 to 1.5 SDs below the mean when compared with their peers with matched age and education level) • Independent in functional abilities • No dementia

Publication or Study Trend

There has been a noticeable increase in both the number of studies and the variety of multimodal intervention combinations from 2014 to 2024 (Figure 2 [15,16,32-78]). Specifically, studies incorporating NIs as part of multimodal MCI interventions grew from 1 in 2016 to 3 in both 2020 and 2023. In addition, the use of technology in delivering multimodal interventions has gradually increased from 1 study in 2015 to 5 in 2023.

Figure 2. The number of studies that incorporated multimodal interventions for mild cognitive impairment throughout the years. CS: current stimulation; CT: cognitive training; NI: nutritional intervention; PI: psychosocial intervention; PT: physical training; SA: social activities.



Types of Multimodal Interventions

The included studies used combinations of 2, 3, and 4 types of interventions. Although most included studies primarily focused on combining PT and CT, there is an increasing number of studies integrating additional interventions to address the multifaceted nature of MCI. The following section details these specific combinations, the types of control groups used, and the overall outcomes across cognitive domains, such as attention, executive function, global cognition, memory, processing speed, and verbal fluency. Comprehensive outcomes, measurement tools, and targeted cognitive domains of the included studies are summarized in the following sections.

Bimodal Interventions

Overview

Most studies that used combined PT and current stimulation [45,67], or PT and NI [56,58] reported significantly greater improvements in various cognition domains compared to active control groups, while the use of combined CT and PI [61,68,69] reported significantly greater improvements compared to inactive control groups. However, studies with combined PT and CT showed mixed results. Conversely, most studies comparing combined PT and PI [39,41,48], CT and electroacupuncture [42], or CT and current stimulation [44,47,53,60,62] with an active or inactive control intervention found little to no significant advantage in the intervention groups (Table 2).

Table 2. Studies that incorporated a bimodal intervention and outcomes (n=35).

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]				
Callisaya et al [57], 2021	44 (49)	• IN[c]	• GC[d] (COWAT[e], DSCT[f], HVLT[g], TMT-A[h], TMT-B[i], and SCWT[j])	• No significant difference between the two groups for all tests (*P*>.05)
Fairchild et al [51], 2024	36 (36)	• IN	• EF[k] (DST[l], TMT-B, and SCWT) • ME[m] (WMS-LM I[n] and II[o], and RAVLT[p]) • PS[q] (TMT-A and SCWT)	• No significant difference between the two groups for all tests (*P*>.05)
Fiatarone Singh et al [38], 2014	27 (22; 24; 27)	• CT only • PT only • PC[r]	• ATT[s] (SDMT[t]) • GC[u] (ADAS-Cog[v]) • EF (Matrices and Similarities subtests of the WAIS-III[w]) • ME (ADAS-Cog, BVRT[x], and LM[y] from WMS-III[z]) • PS (SDMT) • VF[aa] (COWAT and animal naming)	• The intervention group did not show significant differences compared to other groups across all tests (*P*>.05)
Gill et al [59], 2016	23 (21)	• PT only	• GC (composite score) • EF (TMT-A and TMT-B) • ME (AVLT[ab]) • PS (DSST[ac]) • VF (VFT[ad])	• GC (mean difference 0.2, *P*=.04[ae]) • EF (mean difference 0.11, *P*=.60) • ME (mean difference 0.3, *P*=.02) • PS (mean difference −0.06, *P*=.78) • VF (mean difference 0.62, *P*=.003)
Griffiths et al [52], 2020	35 (35)	• IN	• ATT (TMT-A and TMT-B) • EF (BD[af] of the WAIS-IV[ag]) • ME (DST-F[ah], DST-B[ai], DSS[aj] from WAIS-IV, VFT, and WLL[ak])	• ATT (TMT-A: mean difference −27.86, *P*=.36) • EF (mean difference 0.34, *P*=.58) • ME (DST-F and DST-B: mean difference 1.23, *P*=.60; DSS: mean difference 1.030, *P*=.31; VFT-Letter[al]: mean difference 4.59, *P*=.001; VFT-Category[am]: mean difference 2.81, *P*=.23)
Hagovská et al [35], 2016	40 (40)	• PT only	• ATT (SCWT) • GC (ACE[an] and MMSE[ao]) • ME (ACE and AVLT) • PS (DRT-II[ap]) • VF (ACE)	• ATT (SCWT: η2=0.0001, *P*=.97) • GC (ACE: Cohen d=0.71, *P*=.002; MMSE: η2=0.189, *P*=.001) • ME (ACE: Cohen d=0.64, *P*=.007; AVLT: η2=0.173, *P*=.001) • PS (DRT-II: η2=0.033, *P*=.11) • VF (Cohen d=0.73, *P*=.001)
Jeong et al [40], 2021	13 (13)	• HE[aq]	• ATT (TMT-A and TMT-B) • GC (ADAS-Cog and KMMSE[ar]) • PS (DSST)	• Group×time interaction • ATT—TMT-A: *P*<.05; TMT-B: *P*=.01 • GC (ADAS-Cog: *P*=.11); KMMSE (*P*=.72) • PS (*P*=.02)
Kim and Park [63], 2023	21 (21)	• CT only	• EF (EFPT-K[as] and FAB[at])	• EF (EFPT-K: η2=0.132, *P*<.01; FAB: η2=0.305, *P*<.001)
Lam et al [74], 2015	93 (115; 114; 101)	• CT only • PT only • SA only	• ATT (VFT-Category) • GC (ADAS-Cog, CDR-SOB[au], and CMMSE[av]) • ME (list learning delayed recall test)	• ATT (VFT-C: χ2=23.38, *P*<.001) • GC (ADAS-Cog: χ2=3.31, *P*=.2; CDR-SOB: χ2=1.82, *P*=.61; CMMSE: χ2=4.28, *P*=.23) • ME (χ2=3.31, *P*=.35)

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Li et al [66], 2023	19 (35; 30)	• CT only • IN	• ATT, EF, ME, language, and spatial ability (AVLT, STT-A[aw] and STT-B[ax], CFT[ay], SCWT, and BNT-30[az]) • GC (ADAS-Cog and CMMSE)	• ATT, EF, and ME: (AVLT immediate recall: mean difference 0.46, *P*>.05; STT-A: mean difference −0.28, *P*>.05; STT-B: mean difference −0.15, *P*>.05; CFT: mean difference 0.34, *P*>.05; BNT-30: mean difference −1.14, *P*>.05) • GC: ADAS-Cog (mean difference 0.03, *P*>.05; MMSE: mean difference 0.22, *P*>.05)
Liao et al [36,37], 2019, 2020	16 (18)	• PT and CT	• GC (MoCA[ba]) • EF (EXIT-25[bb] and SCWT) • ME (CVVLT[bc])	• Group×time interaction • GC (*P*=.18) • EF (EXIT-25: *P*=.72; SCWT: number: *P*=.84; time: *P*=.32) • ME (CVVLT: immediate recall: *P*=.15; delayed recall: *P*=.12)
Parial et al [15], 2022	25 (26)	• HE	• GC (MoCA) • EF (TMT-B) • ME (DST)	• GC (Wald χ2=26.88, *P*<.001) • EF (Wald χ2=18.67, *P*<.001) • ME (immediate recall: Wald χ2=16.97, *P*<.001; delayed recall: Wald χ2=11.89, *P*<.003)
Park et al [72], 2019	25 (24)	• IN	• GC (ADAS-Cog) • EF (DSST) • ME (DST)	• GC: ADAS-Cog—*P*<.01; 95% CI −1.9 (−0.9 to −2.3) • EF: DSST—*P*<.01; 95% CI −6.3 (−4.2 to −8.4) • ME: DST—*P*=.02; 95% CI −0.6 (−0.3 to 1.6)
Rezola-Pardo et al [70], 2019	43 (42)	• PT only	• Overall cognitive performance (MoCA, symbol search and coding test from WAIS-IV, semantic fluency test, VFT, RAVLT, and TMT-A)	• MoCA (Cohen d=0.15; *P*=.90) • Symbol search (Cohen d=0.06; *P*=.85) • Semantic fluency (Cohen d=0.01; *P*=.28) • VFT (Cohen d=0.02; *P*=.67) • RAVLT (Cohen d=0.11; *P*=.70) • TMT-A (Cohen d=0.04; *P*=.65)
Shimada et al [71], 2018	137 (129)	• HE	• GC (MMSE) • EF (TMT-A and TMT-B) • ME (WMS-LM II and RAVLT) • VF (VFT)	• CG (mean difference 0.8, *P*=.01) • EF (mean difference −0.4, *P*=.35) • ME (WMS-LM II: mean difference 1.0, *P*=.004; RAVLT: mean difference 0.2, *P*=.35) • VF: (mean difference 2.2, *P*=.002)
Styliadis et al [16], 2015	14 (14; 14; 14; 14)	• PT only • CT only • HE • IN	• GC (MMSE)	• The intervention group did not show significant differences compared to other groups across all tests (*P*>.05)
Tao et al [73], 2023	51 (52)	• IN	• GC (MoCA and MMSE)	• GC (MoCA: η2=0.442, *P*=.001; F=39.550; MMSE: η2=0.33, *P*<.001; F=24.614)
Xu et al [65], 2020	5 (6; 6)	• NI only • IN	• GC (HK-MoCA[bd])	• GC: improvement only in the NI[be]-only group (*P*=.008)
PT+PI[bf]				

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Fogarty et al [39], 2016	22 (18)	• PI only	• ATT (TEA[bg]) • ME (RBMT[bh]) • PS (digit symbol test)	• ATT: TEA (visual selective ATT: Cohen d=0.15, *P*=.82; F2,70=0.196; sustained ATT: Cohen d=0.127, *P*=.88; F2,62=0.125; ATT switching: Cohen d=0.0487, *P*=.13; F2,70=2.071) • ME (Cohen d=0.583, *P*=.06; F2,70=2.972) • PS (Cohen d=0.16, *P*=.80; F2,72=0.230)
Kamegaya et al [41], 2014	19 (24)	• IN	• GC (Five-Cog test) • EF (DSST and YKSSTbi)	• GC (F=2.999, *P*=.09; F1,38=2.999) • EF (DSST: F=1.165, *P*=.29; YKSST: F=0.096, *P*=.76; F1,38=1.165)
Shimizu et al [48], 2018	30 (9)	• PT only	• GC (FAB)	• GC: effect size (r=0.002, *P*=.81)
PT+NI				
Blumenthal et al [56], 2019	40 (41; 41; 38)	• PT only • NI only • HE	• GC (modified CDR-SOB) • EF (TMT-A, TMT-B, SCWT, DST-F, DST-B, DSST, Ruff 2 and 7 Test, and animal naming) • ME (HVLT-revised and CFT) • VF (COWAT and animal naming)	• GC (Cohen d=0.36, *P*=.03) • EF (Cohen d=0.4, *P*=.01) • ME (PT factor: Cohen d=0.19, *P*=.24; NI factor: Cohen d=0.15, *P*=.35) • VF (PT factor: Cohen d=0.12, *P*=.45; NI factor: Cohen d=0.23, *P*=.24)
Chobe et al [58], 2022	24 (25; 23)	• PT only • NI only	• ATT (DST-F and DST-B) • EF (TMT-A and TMT-B) • ME (RAVLT) • VF (COWAT)	• ATT (DST-F: F=2.921, *P*=.06; DST-B: F=5.766, *P*=.005) • EF (TMT-A: F=0.837, *P*=.44; TMT-B: F=0.677, *P*=.51) • ME (F=4.727, *P*=.01) • VF (F=3.028, *P*=.06)
CT+electroacupuncture				
Kim et al [42], 2020	16 (16)	• CT only	• GC (ADAS-Cog-K[bi] and MoCA-K[bj])	• GC: (ADAS-Cog-K: Z=−0.38, *P*=.70; MoCA-K: Z=−0.72, *P*=.47)
CT+PI				
Stuerz et al [68], 2022	26 (24)	• Reverse sequential control	• ATT (ACT) • GC (MMSE)	• ATT: ACT (η2<0.01, *P*=.21) • GC: MMSE (η2=0.05, *P*=.09)
Jesus et al [68], 2023a	27 (24)	• IN	• ATT (digit symbol coding of WAIS-III) • GC (ACE-R[bk]) • ME (ACE-R; WL[bl] and LM of WMS-III)	• ATT (F1,49=11.64; η2=0.19, *P*=.001) • GC (F1,49=4.70; η2=0.09, *P*=.04) • ME: ACE-R (F1,49=7.01; η2=0.13, *P*=.01); WL (F1,49=23.76; η2=0.33, *P*<.001); LM (F1,49=10.98; η2=0.18, *P*=.002)
Jesus et al [69], 2023b	98 (101)	• IN	• ATT (digit symbol coding of WAIS-III) • GC (ACE-R) • ME (LM of WMS-III)	• ATT (Cohen d=0.41, *P*<.001) • GC (Cohen d=0.33, *P*<.001) • ME (immediate recall: Cohen d=0.42, *P*<.001; delayed recall: Cohen d=0.35, *P*<.001)
CT+current stimulation				

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Gonzalez et al [60], 2021	22 (24; 21)	• Sham tDCS[bm]+CT • CT only	• ATT (TMT-A) • GC (MoCA) • ME (DST and RBMT, 3rd edition)	• ATT: TMT-A (F=0.64, *P*=.62) • GC: MoCA (F=0.34, *P*=.85) • ME: DST (F=1.99, *P*=.09); RBMT-3 (F=0.13, *P*=.96)
Jones et al [62], 2023	13 (12)	• PC+CT	• ATT (ACE-X[bn]) • ME (ACE-X)	• ATT (Cohen d=0.94, *P*=.047) • ME (η2=0.05, *P*=.35)
Martin et al [53], 2019	33 (35)	• CT+PC	• ME (PALT[bo] and RVIP[bp]) • PS (SDMT) • Subjective cognitive functioning (CFQ[bq])	• ME: PALT (F=0.26, *P*=.77); RVIP (F=0.2, *P*=.82) • PS (F=1.39, *P*=.25) • Subjective cognitive functioning (F=1.25, *P*=.29)
Lau et al [44], 2024	11 (11)	• CT+PC	• GC (MoCA) • EF (TMT-A, TMT-B, and Tower of London) • ME (N-back task and CVVLT)	• GC: MoCA (η2=0.05, *P*=.35) • EF: TMT-A (η2=0.18, *P*=.06); TMT-B (η2=0.12, *P*=.13) • Tower of London (η2=0.03, *P*=.42); • ME: N-back (1-back: η2=0.17, *P*=.07; 2-back: η2=0.16, *P*=.07); CVVLT (verbal memory: η2=0.12, *P*=.07; delayed recall: η2=0.02, *P*=.59)
Senczyszyn et al [47], 2023	13 (13; 12)	• Current stimulation only • Sham current stimulation only	• GC (CANTAB[br]) • VF (Verbal Fluency FAS test)	• GC: CANTAB (Swms6: η2=0.042, *P*>.05; Palta4: η2=0.019, *P*=.03; Prmpci: η2=0.091, *P*=.02) • VF (η2=0.042, *P*>.05)
PT+current stimulation				

Study	Sample size (control or controls sample size)	Control or controls	Measured cognitive domains	Clinical outcomes
Liao et al [45], 2021	10 (10)	• PT+PC	• GC (MoCA) • EF (TMT-A and TMT-B) • ME (change detection task and CVVLT)	• GC: F1,18=0.246; η2=0.246, *P*=.63 • EF: TMT-A (F1,18=1.022; η2=0.054, *P*=.33); TMT-B (η2=0.271, *P*=.02) • ME: change detection task (F1,18=1.046; η2=0.058, *P*=.32); CVVLT (F1,18=0.024; η2=0.001, *P*=.88)
Xu et al [67], 2023	44 (49; 44; 43)	• PT+sham current stimulation • Sham PT+current stimulation; sham PT+sham current stimulation	• ATT (TAP[bs]) • GC (MoCA) • EF (SCWT) • ME (AVLT, ROCF, and Chinese Wechsler Memory Scale—Revised MQ[bt])	• ATT (auditory: F=2.66, *P*=.05; visual: F=4.536, *P*=.004; sustained reaction: F=3.609, *P*=.02) • GC (F=7.415, *P*<.001) • EF (F=.058, *P*=.98) • ME: AVLT (immediate recall: F=3.207, *P*=.03; delayed recall: F=1.13, *P*=.34); ROCF (copy: F=2.489, *P*=.062; recall: F=0.571, *P*=.64); MQ (F=3.584, *P*=.02)

[a]PT: physical training.
[b]CT: cognitive training.
[c]IN: inactive control.
[d]GC: global cognition.
[e]COWAT: Controlled Oral Word Association Test.
[f]DSCT: Digit Symbol Substitution Test.
[g]HVLT: Hopkins Verbal Learning Test.
[h]TMT-A: Trail Making Test–A.
[i]TMT-B: Trail Making Test–B.
[j]SCWT: Stroop Color-Word Test.
[k]EF: executive function.
[l]DST: digit span test.
[m]ME: memory.
[n]WMS-LM I: Wechsler Memory Scale-Revised–Logical Memory I.
[o]WMS-LM II: Wechsler Memory Scale-Revised–Logical Memory II.
[p]RAVLT: Rey Auditory Verbal Learning Test.
[q]PS: processing speed.
[r]PC: placebo control.
[s]ATT: attention.
[t]SDMT: Symbol Digit Modalities Test.
[u]GC: global cognition.
[v]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.
[w]WAIS-III: Wechsler Adult Intelligence Scale–III.
[x]BVRT: Benton Visual Retention Test.
[y]LM: logical memory.
[z]WMS-III: Wechsler Memory Scale–III.
[aa]VF: verbal fluency.
[ab]AVLT: Auditory Verbal Learning Test.
[ac]DSST: Digit Symbol Substitution Test.
[ad]VFT: Verbal Fluency Test.
[ae]Italicized values indicate statistical significance.
[af]BD: block design.
[ag]WAIS-IV: Wechsler Adult Intelligence Scale–IV.
[ah]DST-F: Digit Span Test–Forward.
[ai]DST-B: Digit Span Test–Backward.

[aj]DSS: digit span sequence.
[ak]WLL: Word-List Learning Test.
[al]VFT-Letter: Verbal Fluency Test–Letter.
[am]VFT-Category: Verbal Fluency Test–Category.
[an]ACE: Addenbrooke Cognitive Examination.
[ao]MMSE: Mini-Mental State Examination.
[ap]DRT-II: Disjunctive Reaction Time.
[aq]HE: health education.
[ar]KMMSE: Korean version of the Mini-Mental State Examination.
[as]EFPT-K: Executive Function Performance Test.
[at]FAB: Frontal Assessment Battery.
[au]CDR-SOB: Clinical Dementia Rating Sum of Boxes.
[av]CMMSE: Mini-Mental State Examination, Chinese version.
[aw]STT-A: Shape Trail Test–A.
[ax]STT-B: Shape Trail Test–B.
[ay]CFT: Complex Figure Test.
[az]BNT-30: Boston Naming Test.
[ba]MoCA: Montreal Cognitive Assessment.
[bb]EXIT-25: Executive Interview-25.
[bc]CVVLT: Chinese version of the Verbal Learning Test.
[bd]HK-MoCA: Montreal Cognitive Assessment, Hong Kong version.
[be]NI: nutritional intervention.
[bf]PI: psychosocial intervention.
[bg]TEA: Test of Everyday Attention.
[bh]RBMT: Rivermead Behavioral Memory Test.
[bi]ADAS-Cog-K: Alzheimer Disease Assessment Scale–Cognitive, Korean version.
[bj]MoCA-K: Montreal Cognitive Assessment, Korean version.
[bk]ACE-R: Addenbrooke Cognitive Examination–Revised.
[bl]WL: waitlist control.
[bm]tDCS: transcranial direct current stimulation.
[bn]ACE-X: Adaptive Cognitive Evaluation-Explorer
[bo]PALT: Paired Associative Learning Test.
[bp]RVIP: Rapid Visual Information Processing.
[bq]CFQ: Cognitive Failures Questionnaire.
[br]CANTAB: Cambridge Neuropsychological Test Automated Battery.
[bs]TAP: Test Of Attentional Performance.
[bt]MQ: memory quotient.

The Combination of PT and CT

Among these 18 studies, 4 used a single-modal active control group [34,59,63,70], 3 compared against a health education group [15,40,71], and 5 used a passive control group [51,52,57,72,73]. Two studies used both a single-modal active control and a passive control group [65,66]. One study compared results with 2 single-modal active controls and a placebo group [38], while another used 3 single-modal active control groups [74]. In addition, 1 study contrasted results with 2 single-modal active controls, a health education group, and a passive control group [16]. Another study used a bimodal active control group [36].

Among these studies, 7 reported significantly greater improvements in all measured cognitive functions within the bimodal intervention groups [15,35,40,52,59,63,72]. Three studies revealed improvements in both bimodal intervention and control groups [16,37,61,70]. One study found greater improvements in executive function, global cognition, and verbal fluency within the PT-only control group, whereas attention and processing speed improved significantly across all groups [38]. Another study reported no changes in cognitive function across any group [57]. In addition, 1 study observed improvements in attention, episodic memory, and working memory across all groups but no changes in global cognition [74]. Finally, 1 study reported greater improvement in global cognition within the NI-only control group [65].

Three Studies Investigated the Combination of PT and PI

Among these 3 studies, 2 studies compared results with a single-modal active control group [39,48], and one compared with a passive control group [41]. One study found that both the bimodal intervention and control groups enhanced attention, memory, and processing speed, but only the intervention group significantly improved executive function [39]. Another study

revealed improvements in global cognition only within the bimodal intervention group, although both groups improved in executive function [41]. In addition, one study reported significantly greater improvements in global cognition exclusively in the bimodal intervention group [48].

Two Studies Investigated the Combination of PT and NI

One study compared the intervention group with 2 single-modal active control groups and a health education control group [56]. It found that both the bimodal intervention and PT-only groups showed greater improvements in executive function and global cognition, although no gains were observed in memory or verbal fluency across groups [56]. Conversely, another study contrasted the intervention group with 2 single-modal active control groups [58] and revealed improvements in global cognition in all groups [58].

One Study Investigated the Combination of CT and Electroacupuncture

This study compared the intervention group with a single-modal active control group. It found that all groups demonstrated significant improvements in global cognition [42].

Three Studies Investigated the Combination of CT and PI

One study compared a reverse sequential control group [68], whereas 2 studies compared results with a passive control group [61,69]. Results from the latter 2 studies revealed significantly greater improvements in attention, episodic memory, and global cognition in the bimodal intervention group [61,69]. Conversely, the study with a reverse sequential control group found enhanced global cognition in both groups. Notably, the group starting with the PI showed more significant attention improvements after the initial training, while the group beginning with CT exhibited greater improvements following the second, combined training phase [68].

Five Studies Investigated the Combination of CT and Current Stimulation

Two of these studies compared the intervention group with a bimodal active control group [53,62]. One study contrasted the intervention group with a single-modal and a bimodal active control group [60]. Another study used a combination of CT and sham current stimulation as a control group [44]. One study included a current stimulation–only group, a sham current stimulation–only group, and a sham current stimulation control group [47].

One study observed improvements in everyday memory and global cognition across all groups, with attention improvements unique to the bimodal intervention group [60]. However, the CT-only control group showed greater working memory improvements than the bimodal intervention group [60]. Notably, transcranial direct current stimulation (tDCS) appeared to enhance the efficacy of CT in processing speed [60], despite not being a direct target. Another study showed that the bimodal intervention group demonstrated more significant attention improvements than both the placebo and CT groups [62]. In addition, 1 study found improvements in attention; processing speed; subjective cognitive functioning; and verbal, visual, and working memory across all groups [53]. Another study revealed improvements in episodic memory, executive function, and visual memory in both bimodal and active control groups [44].

Two Studies Investigated the Combination of PT and Transcranial Brain Stimulation

One study compared the intervention group with a bimodal active control group (PT and sham brain stimulation) [45], while another study used 2 bimodal active controls (PT and sham stimulation and sham PT with brain stimulation) along with a passive control group [67].

Results showed improvements in episodic memory, executive function, global cognition, and visual working memory in both the bimodal intervention and control groups [45]. However, the bimodal intervention group displayed greater enhancements across all cognitive functions compared to the control groups [67].

Trimodal Interventions

Overview

Studies comparing combined PT, CT, and SA with health education controls reported significantly better cognitive improvements, especially in the memory domain [55,75,76]. Similarly, 3 out of 5 studies using combined PT, CT, and NI [46,49,54] showed significantly greater improvements in global cognition compared to control groups (Table 3).

Table 3. Studies that incorporated trimodal intervention and outcomes (k=8).

Study	Sample size (control sample size)	Control	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]+NI[c]				
Bray et al [50], 2023	18 (15; 20; 15; 14)	• PT+CT+PC[d] • PT+NI+PC • PT+PC+PC • PC only	• GC[e] (ADAS-Cog[f]) • EF[g] and ME[h] (ADAS-Cog, TMT-A[i], and TMT-B[j])	• The intervention group did not show significant differences compared to other groups across all tests (*P*>.05)
Kobe et al [43], 2016	13 (9)	• NI[k]	• GC (AVLT[l], DST[m], VFT[n], TMT-A, TMT-B, and SCWT[o])	• No significant difference between the two groups for all tests (*P*>.05)
Montero-Odasso et al [54], 2023	34 (35; 37; 35)	• PT+CT+sham NI • PT+sham CT+NI • PT+sham CT+sham NI • Sham PT+sham CT+sham NI	• GC (ADAS-Cog) • EF (ADAS-Cog Plus variant)	• GC (mean difference −2.64, *P*=.005[p], Cohen d=0.71) • EF (ADAS-Cog Plus variant: no significant improvement; *P*>.05)
Phoemsapthawee et al [46], 2022	20 (20; 18)	• PT only • PC	• GC (MMSE[q]) • EF (TMT-A and TMT-B) • ME (DST-F[r] and DST-B[s])	• GC (F=13.158; η2=0.328, *P*<.001) • EF: TMT-A (F=30.142; η2=0.527, *P*<.001); TMT-B (η2=0.376, *P*<.001) • ME: DST-F (F=17.208; η2=0.389, *P*<.001); DST-B (η2=.495, *P*<.001)
Yang et al [49], 2022	55 (57)	• IN	• GC (MoCA[t])	• GC: group×time interaction ($\chi^2$3=303.928; *P*<.001)
PT+CT+SA				
Bae et al [55], 2019	41 (42)	• HE[u]	• GC (MMSE) • EF (TMT-A and TMT-B) • ME (spatial span task)	• GC (*P*=.14) • EF: TMT-A (*P*=.43); TMT-B (*P*=.68) • ME (*P*=.02)
Lee et al [75], 2023	140 (140)	• HE	• ATT[v] and EF (TMT-A and TMT-B) • ME (NCGG-FAT[w]) • PS[x] (DSST[y])	• ATT and EF: TMT-A (F=0.25); TMT-B (F=0.00) • ME (F=5.04, *P*<.05) • PS (F=0.67)
Lin et al [76], 2020	61 (61)	• HE	• ATT (SLUMS[z]) • GC (SLUMS) • EF (SLUMS)	• ATT (Wald χ2=50.84, *P*<.001) • GC (Wald χ2=252.81, *P*<.001) • EF (Wald χ2=115.99, *P*<.001)

[a]PT: physical training.

[b]CT: cognitive training.

[c]IN: inactive control.

[d]PC: placebo control.

[e]GC: global cognition.

[f]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.

[g]EF: executive function.

[h]ME: memory.

[i]TMT-A: Trail Making Test–A.

[j]TMT-B: Trail Making Test–B.

[k]NI: nutritional intervention.

[l]AVLT: Auditory Verbal Learning Test.

[m]DST: Digit Span Test.

[n]VFT: Verbal Fluency Test.

[o]SCWT: Stroop Color-Word Test.

[p]Italicized values indicate statistical significance.

[q]MMSE: Mini-Mental State Examination.

[r]DST-F: Digit Span Test–Forward.

[s]DST-B: Digit Span Test–Backward.

[t]MoCA: Montreal Cognitive Assessment.

[u]HE: health education.

[v]ATT: attention.

[w]NCGG-FAT: National Center for Geriatrics and Gerontology–Functional Assessment Tool.

[x]PS: processing speed.

[y]DSST: Digit Symbol Substitution Test.

[z]SLUMS: Saint Louis University Mental Status Examination.

Five Studies Investigated the Combination of PT, CT, and NI

One study compared the trimodal intervention group with an NI-only active control group [43], another compared it with both a PT-only active control and a placebo group [46], and a third study used an inactive control group [49]. One study contrasted outcomes with 4 control groups: combined PT and CT with placebo NI; combined PT and NI with sham CT; combined PT with sham CT and sham NI; and an inactive control group [54]. The fifth study compared the trimodal intervention with various combinations of PT, CT, and NI placebos [50].

In the first study, the trimodal intervention group study reported no substantial changes in cognitive functions across all groups [43]. Conversely, the second study found the intervention group showed significantly greater improvements in all cognitive functions compared with the placebo group, but not when compared to the PT-only control group [46] and the third study observed more significant enhancements in global cognition in the trimodal intervention group [49]. The fourth study also found greater improvements in global cognition in the trimodal intervention group, but no significant changes in other measures [54]. In the last study, the trimodal intervention group showed no significant difference compared to the controls [50].

Three Studies Investigated the Combination of PT, CT, and SA

All 3 studies contrasted the intervention group with a health education control group, revealing significant improvements in global cognition and memory for the trimodal intervention group [55,76]. In addition, 1 study found significant enhancements in memory compared to control groups [55], while another reported significant improvements in both attention and executive function during a 12-month follow-up in the intervention group [76].

Quadrimodal Interventions

Most studies incorporating quadrimodal interventions reported small to moderate effects across various cognition domains, especially in global cognition (Table 4).

Table 4. Studies that incorporated quadrimodal interventions and outcomes (n=3).

Study	Sample size (control sample size)	Control	Measured cognitive domains	Clinical outcomes
PT[a]+CT[b]+PI[c]+SA[d]				
Maffei et al [77], 2017	55 (58)	IN[e]	• GC[f] (ADAS-Cog[g])	• GC (Effect Size=−0.55, *P*<.001[h])
Straubmeier et al [64], 2017	208 (154)	IN	• GC (MMSE[i])	• GC (Cohen d=0.26, *P*=.01)
PT+CT+PI+NI[j]				
Liu et al [78], 2023	86 (106)	IN	• GC (MMSE and MoCA[k]) • ME (AVLT[l] and PALT[m])	• GC (Hedges g=0.40, *P*=.03) • ME (*P*=.05)

[a]PT: physical training.

[b]CT: cognitive training.

[c]PI: psychosocial intervention.

[d]SA: social activities.

[e]IN: inactive control.

[f]GC: global cognition.

[g]ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive.

[h]Italicized values indicate statistical significance.

[i]MMSE: Mini-Mental State Examination.

[j]NI: nutritional intervention.

[k]MoCA: Montreal Cognitive Assessment.

[l]AVLT: Auditory Verbal Learning Test.

[m]PALT: Paired Associative Learning Test.

Two Studies Investigated the Combination of PT, CT, PI, and SA

The 2 studies compared a quadrimodal intervention group with a passive control group [64,77]. One study found significantly greater improvements in global cognition in the quadrimodal intervention group [77]. In contrast, the other study reported no significant improvement in cognitive function within the intervention group, but the control group experienced deterioration [64].

One Study Investigated the Combination of PT, CT, PI, and NI

Compared to a passive control group, the quadrimodal intervention group showed significantly larger posttreatment improvements in global cognition and memory. However, these effects showed no difference between the two groups at the 12-month follow-up [78].

The Use of Technology in MCI Interventions

Between 2015 and 2024, a total of 16 studies incorporated technology to enhance intervention delivery. The Nintendo Wii (n=1) was first used as a hardware platform for delivering PT [16]. Subsequent studies used various technologies, including virtual reality (VR) technology (n=1) for combined PT and CT intervention [37]; Nintendo Switch (n=1) [44]; and iPad (n=2) [50,57] for sensorimotor, visuomotor [54], and CT [50], respectively.

In addition, 13 studies used a range of computerized programs for CT targeting functions such as attention, memory, and executive function. These included AKL-T01 (n=1) [62], Brain Fitness (n=1) [16], COGPACK (n=2) [53,68], CogniPlus (n=1) [32], NeuronUP (n=1) [60], Neuropeak (n=1) [54], and RehaCom (n=3) [42,47,63] and 2 studies with no specified programs [44,49]. In addition, 1 study used the FitForAll program [16] for PT. Studies that incorporated CCT alongside other modalities showed significantly better improvements, particularly in global cognition, than studies with either active and inactive control groups. Details on tasks, duration, and major findings of these programs are presented in Multimedia Appendix 3.

Five studies incorporated tDCS with CT [44,47,53,60,62], while 2 studies combined tDCS with PT [45,67]. Specifically, they applied 1.5mA to 2 mA of tDCS to the left dorsolateral prefrontal cortex (DLPFC) or 1.5 mA of transcranial alternating current stimulation (tACS) on the prefrontal cortex [62].

Acceptability and User Experience

User compliance is influenced by the acceptability and experience of treatments. Four studies indicated high acceptability and positive user experiences regarding multimodal interventions. Jesus et al [61] reported that 24 out of 27 participants demonstrated a 97% acceptability rate for the treatment effect and frequency of a combined CT and PI. Another study on a combined PT, CT, and PI intervention reported that 78% (39/50) of participants rated the program as “very good,” 20% (10/50) rated it as “good,” and 2% (1/50) rated it as “pleasurable” [68]. In addition, a study on a combined PT, CT, and SA program showed that 86% (35/41) of

participants were satisfied with the intervention duration and 81% (33/41 participants) expressed a desire to continue [55].

Similarly, a study on a combined PT and CT intervention reported that 39 out of 44 (89%) participants perceived subjective benefits. In terms of overall satisfaction, 14 out of 44 (32%) participants rated it as "very good," 24 out of 44 (54%) participants rated it as "good," and 6 out of 44 (13%) participants rated it as "neither good nor bad." In addition,12 out of 44 (27%) participants considered the user-friendliness of the intervention as "very good" [57].

Dosage Effect and Health Economics

While treatment dosage may affect clinical outcomes, only 1 study evaluated the dosage effects of a quadrimodal intervention (combined PT, CT, PI, and SA) [64]. It found that 1 to 2 sessions per week of the quadrimodal intervention in daycare centers yielded no significant difference in cognitive outcomes compared with 3 to 5 sessions per week [64]. As such, the study suggested that less-frequent sessions (1-2 wk) might be as effective as more frequent sessions, thus further studies are warranted to validate these findings and test with other multimodal interventions.

In addition, while the cost-effectiveness of multimodal cognitive interventions is important for clinical practice, none of the included studies examined the cost-effectiveness of the identified intervention.

Discussion

Overview

This scoping review examines the current research landscape on nonpharmacological multimodal interventions for MCI, highlighting publication trends, intervention types, technology use, user experience, and dosage effects. Over the past decade, the variety of multimodal intervention combinations has increased. The most common interventions combine PT and CT, with additional components, such as NI, electroacupuncture, PI, and current stimulation. Cognitive outcome measures are diverse, targeting various domains. Bimodal and trimodal interventions generally outperform single-modal ones in improving global cognition, attention, and executive function. Notably, PT with current stimulation and PT with NI often demonstrate better cognitive improvements compared to active controls, while CT with PI shows more improvements compared to inactive controls. However, most studies on combined PT and CT report mixed results. Quadrimodal interventions also show superior improvements, although their long-term effect remains uncertain. This review, following the Arksey and O'Malley framework, did not conduct a meta-analysis or risk of bias assessments. The included studies suggested that trimodal and quadrimodal interventions, especially those including SA, might offer better cognitive outcomes, but findings should be interpreted cautiously due to the limited number of studies. Technologies, such as VR, gaming, computerized programs, and transcranial stimulation, are increasingly adopted in MCI interventions. Although technology-assisted training showed significant improvements in various cognitive domains comparable to traditional trainings, results should be interpreted with care given the limited studies. Future research should investigate whether technology-assisted training, with or without additional modalities, offers significantly better cognitive improvements than traditional training. Although only 4 included studies assessed the acceptability of multimodal interventions, they were generally well-received by users. Only 1 included study evaluated the dosage effects of a quadrimodal intervention, suggesting that 3 or more sessions per week may not be beneficial, although this finding should be interpreted cautiously. As none of the included studies evaluated the cost-effectiveness of these interventions, future studies should address this gap.

Types of Modalities

Aside from traditional PT like aerobic and stretching exercises, recent research has been expanding to explore various PT types beyond traditional methods, including mind-body exercises such as tai chi [66], yoga [58], and dancing [15] as viable components in multimodal interventions for managing MCI.

Mind-body exercises, which engage the mind to influence bodily functions, are popular among older adults, partly due to cultural preferences in certain populations. The effectiveness of such exercises, such as tai chi, varies between Asian [74,79] and Western populations [39], highlighting cultural influences. tai chi, for instance, demands whole-body coordination, rhythmic movements, dynamic weight shifting, single-limb support, integrating movement recall (memorization and concentration), spatial orientation, and cognitive activities such as attention and executive control. This provides simultaneous PT and CT that could be beneficial for older adults with MCI [20,66]. In addition, these mind-body exercises offer relaxation and social support in group settings, potentially improving mood and motivation to participate in activities. Given their multifaceted benefits, mind-body physical activities may be incorporated into multimodal interventions for MCI.

With the growing interest in multimodal interventions, there is a noticeable increase in incorporating various interventions in addition to combined PT and CT. Modifiable lifestyle factors, including NI, are crucial for cognitive improvements among older adults with MCI. However, the effectiveness of NI as a standalone intervention has yielded mixed results for cognitive outcomes. Different dietary patterns and supplements have been suggested to promote cognitive health. For instance, the Mediterranean-Dietary Approaches to Stop Hypertension Intervention for Neurodegenerative Delay diet is popular. However, a large-scale study found that the Mediterranean-Dietary Approaches to Stop Hypertension Intervention for Neurodegenerative Delay diet was not significantly superior to a control diet with mild caloric restriction over 3 years [80]. In addition, systematic reviews suggest that certain single-nutrient supplements, such as folate, vitamin E, omega-3 fatty acids, and probiotics, show promising but preliminary results, often based on weak evidence or low-quality studies [81-83].

NI has drawn attention to its potential in multimodal interventions for MCI. All the included studies that incorporated NI demonstrated promising outcomes when combined with PT and CT compared to the control groups. However, some studies

lacked detailed descriptions of their use of dietary guidelines, limiting the understanding of whether the specific dietary changes could be contributed to cognitive improvement. The current findings suggest that while a single NI may offer limited results, a combination of different modalities can provide a synergistic protective effect, highlighting the effectiveness of a multimodal approach over a single intervention.

In addition, there is an increasing use of PI and SA as supplementary components. Previous studies suggested that engaging in SA could help delay the development of dementia and cognitive decline [84,85]. SA and leisure activities not only have therapeutic benefits but also allow participants to engage with peers, fostering long-term adherence to interventions. It is important to note that the number of studies that incorporated the use of PI and SA is still relatively small; more studies are warranted to strengthen whether the implementation of additional PI and SA might provide additional benefits in a multimodal intervention for MCI.

Collectively, all the aforementioned intervention components play a role in enhancing mental health and cognitive functions across various aspects of MCI, suggesting that a larger number of combined modalities could be beneficial. This aligns with a previous systematic review that suggested an increased number of modalities might yield greater improvement in cognition [21]. Indeed, most of the included studies incorporating 3 or 4 modalities reported notable improvements in various cognitive domains, especially when additional modalities (including SA, NI, and PI) were combined with PT and CT. However, there is no consensus regarding the optimal number or combinations of modalities for multimodal interventions. Furthermore, it is worth noting that adopting multiple lifestyle changes may pose challenges for older adults. As Schneider and Yvon [86] suggested, such changes are more feasible for older adults in better health and with higher education levels. In current MCI management, while health advice is important, it may not sufficiently inform or motivate individuals with MCI to adopt new management strategies. Therefore, health education emerges as a crucial component for the successful implementation of lifestyle changes because it provides vital information on lifestyle modifications and guides the execution of physical or cognitive exercises at home.

Nonetheless, it is noteworthy that most of the included studies only compared the effectiveness of 3 or more modal combinations with single-modal interventions, health education, or passive control groups. It remains unclear whether there are more modalities or specific combinations that would offer greater benefit to individuals with MCI.

The Current Use of Technology

Beyond traditional interventions, technology is increasingly embraced by researchers for its potential as a valuable asset in delivering multimodal interventions for clinicians and researchers from 2015 (n=1) [16] to 2023 (n=4) [47,54,62,63]. CCT comprises guided drill-and-practice on standardized tasks aimed at various cognitive domains that gradually progress with adjustable difficulty based on user performance. This approach is both cost-effective and safe for widespread use [87].

This review identified several computer programs that provide personalized CCT targeting various cognitive domains. Our findings concur with a previous systematic review, showing CCT and VR-CT as effective in enhancing attention, executive function, global cognition, memory, processing speed, verbal fluency, and visuospatial ability in older adults at risk of cognitive decline [25]. These outcomes are on par with traditional in-person PT and CT, suggesting CCT and VR-CT as promising ways for broad-scale cognitive intervention delivery.

Interactive video game–like programs offer training across multiple cognitive domains, including attention, executive function, processing speed, and visuomotor and visuospatial abilities, providing more engaging and motivating experiences for older adults with MCI than traditional face-to-face CT [26,88-90]. These programs also facilitate PT through diverse hardware. For instance, one included study used Nintendo Wii, Wii Remote, and Wii Balance Board for physical exercises (eg, aerobic exercises) [16]. Advanced technologies such as the latest Nintendo Switch, which tailored physical exercise games and hardware (eg, Ring Fit Adventure), adapt challenges to individual capacities, progressively improving cognitive function in older adults with MCI.

In addition to the computerized training program, the included studies that incorporated VR technology in delivering combined PT and CT demonstrated significant improvement in executive function and global cognition. These improvements were greater and consistent with the improvements from the combined PT and CT control group without the use of VR [36,37]. This positions VR as a promising tool for integrated therapy. As suggested by recent systematic reviews, VR could act as an assistive device to deliver training and help improve executive function and global cognition in individuals with MCI [25,88]. Compared to traditional PT and CT, VR-PT and VR-CT offer accessibility, cost-effectiveness, and immersive personalized experience to facilitate skill transfer to activities of daily living [36,37,90]. With advancements in technology, VR equipment has become more user-friendly, affordable, and suitable for home use and rehabilitation, benefiting older adults with mobility issues. However, potential side effects such as stimulator sickness, discomfort, tiredness, mood change induced by immersive experience, and some health-related issues from wearing the VR headgear [90] highlight the need for clinician oversight to adjust treatment as necessary.

Although the development of computerized training programs can be costly, the potential for widespread deployment can offset the expenses of hiring numerous qualified assessors or clinicians and the costs associated with traveling for face-to-face training. Future research should explore integrating these programs with multimodal interventions, such as NI, to enhance their effectiveness and improve the overall well-being of individuals with MCI.

MCI affects multiple cognitive functions. The DLPFC plays a crucial role in functional connectivity with other brain regions and controls various cognitive functions, including attention, decision-making, planning, and working memory. Research has found that individuals with MCI exhibit functional disconnection

in the left DLPFC, leading to attention and working memory deficits [60]. This review included studies that used noninvasive brain stimulation techniques (eg, tDCS [45,53,60] targeting the left DLPFC and tACS targeting the prefrontal cortex); results from these studies suggested that using tDCS or tACS alongside PT or CT could enhance the effectiveness of PT or CT by improving cognitive capacity and reducing the load needed to perform cognitive tasks, thereby increasing processing speed and potentially serving as a viable component in future studies.

Taken together, technology stands as a valuable adjunct in administrating PT and CT, offering standardized, cost-effective, and assessable personalized experiences. Such technological integration could help researchers develop comprehensive management strategies for the widespread treatment of MCI in older adults.

Acceptability and User Experience

Acceptability and user experience are pivotal for the success of interventions aimed at managing MCI, which directly influence its adherence rates. Adherence, in turn, significantly affects the effectiveness of multimodal interventions. Therefore, the interplay between user experience, acceptability, and adherence is vital in developing and implementing MCI interventions. However, most included studies have not investigated these factors, which may hinder clinicians or senior management in choosing or recommending treatment options. One included study suggested that the low adherence rate in older adults might be attributed to family responsibilities, health issues, and social commitments. Leveraging technology, particularly eHealth and mHealth tools, could improve adherence by alleviating barriers to adherence [65]. These tools could be used to facilitate easier implementation and monitoring of interventions, thereby reducing the time and cost required for broader deployment across various populations and regions [91]. Although none of the included studies have implemented such technological applications, future research should explore the potential of mobile technology in enhancing adherence in older adults with MCI.

Limitations and Future Directions

While this scoping review offers a comprehensive overview of current study directions in multimodal interventions for MCI management, it has several limitations. First, by focusing solely on experimental studies with control groups, it may have overlooked relevant insights from qualitative research and gray literature. Second, this scoping review only included English articles, which may introduce cultural bias. Third, this review concentrated exclusively on the cognitive outcomes, omitting physical and psychological outcomes that might be crucial considerations for policy or clinical decisions. Fourth, several included studies had relatively small sample sizes, ranging from 19 to 27 participants [16,43,45,62,65]. Therefore, results should be interpreted with caution. Fifth, adhering to the Arksey and O'Malley framework and PRISMA-ScR guidelines meant no risk of bias assessments or evidence syntheses were conducted. Future systematic reviews or network meta-analyses should aim to summarize evidence regarding the relative effectiveness of different combinations of various multimodal interventions for MCI management.

While the current American Academy of Neurology guidelines for MCI management only recommend PT or CT, recent studies highlight the potential benefits of NI, PI, and SI for MCI management. With the annual conversion rate of MCI to dementia ranging from 10% to 15% per year [7], further research should explore various multimodal intervention combinations, focusing on follow-up, treatment duration, and frequency to potentially reduce or delay conversion rates. Although some included studies used brain imaging, most relied on self-reported or assessment scores, which might not accurately reflect real-world improvements. Future research should incorporate brain imaging technologies to validate results more robustly and elucidate the mechanisms underlying clinical improvements in cognitive function.

Considering the limited research on the cost-effectiveness of multimodal interventions for older adults with MCI and the growing aging population, future studies should investigate the economic viability of these interventions. In addition, the unclear dosage effects of different multimodal intervention combinations warrant investigation. Only one included study revealed that attending 1 or 2 sessions per week of a quadrimodal intervention was not significantly different from 3 to 5 sessions per week, suggesting that fewer sessions of quadrimodal intervention (PT, CT, PI, and SA) might still significantly improve global cognition [64]. However, this finding from a single study is not conclusive, further research is needed to determine the optimal dosage and combination of modalities to optimize the effectiveness of multimodal interventions for MCI.

Previous studies highlight the potential of lifestyle modifications such as dietary patterns in managing cognitive decline in older adults [92-94]. However, these studies used diverse nutrition supplements or dietary patterns. Similarly, while social engagement could help manage MCI [5], there is a dearth of studies on its effectiveness in enhancing physical and emotional well-being in older adults with MCI.

Finally, the included studies used diverse diagnostic criteria and cognitive measurements with different cutoff scores, potentially resulting in heterogeneous participant cohorts that could affect the results. An international consortium should be formed to develop standardized outcome measures for consistent outcome comparisons.

Conclusions

This scoping review provides a comprehensive update on the use of multimodal interventions for improving cognitive functions in older adults with MCI. It presents study directions, multimodal intervention types, respective findings in the included studies, the role of technology in these interventions, and potential research directions. The most common multimodal intervention combines PT and CT. However, various types of interventions such as NI, PI, SI, and brain stimulation have also been incorporated into multimodal interventions more recently. Given the capacity of these interventions to stimulate multiple cognitive domains, the effectiveness of various combinations of modalities should be explored. The research gaps highlighted in this review pave the path for future large-scale clinical trials to help develop more effective management strategies for cognitive decline in older adults with MCI.

Acknowledgments

The authors would like to thank the authors and participants of the included and referenced studies in this scoping review. In addition, they would like to thank the Health and Social Science faculty librarian at the Hong Kong Polytechnic University for her help with our search terms. This scoping review was supported by funding from the Research Institute for Smart Ageing at The Hong Kong Polytechnic University (grant 1-CD4Z). David Ho-Keung Shum is supported by the Yeung Tsang Wing Yee and Tsang Wing Hing Endowed Professorship.

Authors' Contributions

Conceptualization: RCFC (lead), JHSZ (equal), AYLW (equal), YC (supporting), KL (supporting), DHKS (supporting).
Data Curation: RCFC (lead), JHSZ (equal).
Funding Acquisition: DHKS (lead), AYLW (equal).
Methodology: RCFC (lead), JHSZ (equal).
Supervision: AYLW (lead), YC (supporting), KL (supporting), PHFN (supporting), DHKS (supporting).
Visualization: PHFN.
Writing – original draft: RCFC (lead), JHSZ (equal).
Writing – review & editing: AYLC (lead), YC (supporting), KL (supporting), PHFC (supporting), DHKS (supporting).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search keywords and terms used in the current scoping review and results from each database.
[DOCX File , 19 KB - aging_v8i1e70291_app1.docx]

Multimedia Appendix 2
Detailed table of the included studies' and participants' demographic characteristics.
[DOCX File , 76 KB - aging_v8i1e70291_app2.docx]

Multimedia Appendix 3
Table of computerized training programs used in the included studies.
[DOCX File , 25 KB - aging_v8i1e70291_app3.docx]

Multimedia Appendix 4
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Review) Checklist.
[DOCX File , 27 KB - aging_v8i1e70291_app4.docx]

References

1. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
2. Mrakic-Sposta S, Di Santo SG, Franchini F, Arlati S, Zangiacomi A, Greci L, et al. Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in MCI patients: a pilot study. Front Aging Neurosci 2018;10:282 [FREE Full text] [doi: 10.3389/fnagi.2018.00282] [Medline: 30327596]
3. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund L, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. J Intern Med 2004 Sep;256(3):240-246 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
4. Petersen RC. Mild cognitive impairment. Continuum (Minneap Minn) 2016 Apr;22(2 Dementia):404-418 [FREE Full text] [doi: 10.1212/CON.0000000000000313] [Medline: 27042901]
5. Lydon EA, Nguyen LT, Nie Q, Rogers WA, Mudar RA. An integrative framework to guide social engagement interventions and technology design for persons with mild cognitive impairment. Front Public Health 2022;9:750340 [FREE Full text] [doi: 10.3389/fpubh.2021.750340] [Medline: 35096730]
6. Sachdev PS, Lipnicki DM, Kochan NA, Crawford JD, Thalamuthu A, Andrews G, Cohort Studies of Memory in an International Consortium (COSMIC). The prevalence of mild cognitive impairment in diverse geographical and ethnocultural

regions: the COSMIC collaboration. PLoS One 2015;10(11):e0142388 [FREE Full text] [doi: 10.1371/journal.pone.0142388] [Medline: 26539987]
7. Pal K, Mukadam N, Petersen I, Cooper C. Mild cognitive impairment and progression to dementia in people with diabetes, prediabetes and metabolic syndrome: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2018 Nov;53(11):1149-1160 [FREE Full text] [doi: 10.1007/s00127-018-1581-3] [Medline: 30182156]
8. Global action plan on the public health response to dementia 2017-2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017?2025 [accessed 2024-04-29]
9. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-04-29]
10. He CY, Zhou Z, Kan MM, Chan DH, Wong AC, Mok KH, et al. Modifiable risk factors for mild cognitive impairment among cognitively normal community-dwelling older adults: a systematic review and meta-analysis. Ageing Res Rev 2024 Aug;99:102350. [doi: 10.1016/j.arr.2024.102350] [Medline: 38942197]
11. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
12. Vasenina EE, Levin OS, Sonin AG. Sovremennye tendentsii v épidemiologii dementsii i vedenii patsientov s kognitivnymi narusheniiami. Zh Nevrol Psikhiatr Im S S Korsakova 2017;117(6. Vyp. 2):87-95. [doi: 10.17116/jnevro20171176287-95] [Medline: 28980619]
13. Schmitter-Edgecombe M, Dyck DG. Cognitive rehabilitation multi-family group intervention for individuals with mild cognitive impairment and their care-partners. J Int Neuropsychol Soc 2014 Oct;20(9):897-908. [doi: 10.1017/S1355617714000782] [Medline: 25222630]
14. Biazus-Sehn LF, Schuch FB, Firth J, Stigger FD. Effects of physical exercise on cognitive function of older adults with mild cognitive impairment: a systematic review and meta-analysis. Arch Gerontol Geriatr 2020;89:104048. [doi: 10.1016/j.archger.2020.104048] [Medline: 32460123]
15. Parial LL, Kor PP, Sumile EF, Leung AY. Dual-task Zumba gold for improving the cognition of people with mild cognitive impairment: a pilot randomized controlled trial. Gerontologist 2023 Aug 24;63(7):1248-1261. [doi: 10.1093/geront/gnac081] [Medline: 35679826]
16. Styliadis C, Kartsidis P, Paraskevopoulos E, Ioannides AA, Bamidis PD. Neuroplastic effects of combined computerized physical and cognitive training in elderly individuals at risk for dementia: an eLORETA controlled study on resting states. Neural Plast 2015;2015:172192 [FREE Full text] [doi: 10.1155/2015/172192] [Medline: 25945260]
17. Cheng ST. Cognitive reserve and the prevention of dementia: the role of physical and cognitive activities. Curr Psychiatry Rep 2016 Sep 1;18(9):85 [FREE Full text] [doi: 10.1007/s11920-016-0721-2] [Medline: 27481112]
18. Simon SS, Yokomizo JE, Bottino CM. Cognitive intervention in amnestic mild cognitive impairment: a systematic review. Neurosci Biobehav Rev 2012 Apr;36(4):1163-1178. [doi: 10.1016/j.neubiorev.2012.01.007] [Medline: 22322184]
19. Karp A, Paillard-Borg S, Wang HX, Silverstein M, Winblad B, Fratiglioni L. Mental, physical and social components in leisure activities equally contribute to decrease dementia risk. Dement Geriatr Cogn Disord 2006;21(2):65-73. [doi: 10.1159/000089919] [Medline: 16319455]
20. Lissek V, Suchan B. Preventing dementia? Interventional approaches in mild cognitive impairment. Neurosci Biobehav Rev 2021 Mar;122:143-164. [doi: 10.1016/j.neubiorev.2020.12.022] [Medline: 33440197]
21. Salzman T, Sarquis-Adamson Y, Son S, Montero-Odasso M, Fraser S. Associations of multidomain interventions with improvements in cognition in mild cognitive impairment: a systematic review and meta-analysis. JAMA Netw Open 2022 May 02;5(5):e226744 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.6744] [Medline: 35503222]
22. Yang C, Moore A, Mpofu E, Dorstyn D, Li Q, Yin C. Effectiveness of combined cognitive and physical interventions to enhance functioning in older adults with mild cognitive impairment: a systematic review of randomized controlled trials. Gerontologist 2020 Nov 23;60(8):633-642. [doi: 10.1093/geront/gnz149] [Medline: 31697831]
23. Sohn M, Yang J, Sohn J, Lee JH. Digital healthcare for dementia and cognitive impairment: a scoping review. Int J Nurs Stud 2023 Apr;140:104413. [doi: 10.1016/j.ijnurstu.2022.104413] [Medline: 36821951]
24. Di Lorito C, Bosco A, Rai H, Craven M, McNally D, Todd C, et al. A systematic literature review and meta-analysis on digital health interventions for people living with dementia and mild cognitive impairment. Int J Geriatr Psychiatry 2022 Jun;37(6):5730 [FREE Full text] [doi: 10.1002/gps.5730] [Medline: 35588315]
25. Coyle H, Traynor V, Solowij N. Computerized and virtual reality cognitive training for individuals at high risk of cognitive decline: systematic review of the literature. Am J Geriatr Psychiatry 2015 Apr;23(4):335-359. [doi: 10.1016/j.jagp.2014.04.009] [Medline: 24998488]
26. Zhang H, Huntley J, Bhome R, Holmes B, Cahill J, Gould RL, et al. Effect of computerised cognitive training on cognitive outcomes in mild cognitive impairment: a systematic review and meta-analysis. BMJ Open 2019 Aug 18;9(8):e027062 [FREE Full text] [doi: 10.1136/bmjopen-2018-027062] [Medline: 31427316]

27. Shao Z, Hu M, Zhang D, Zeng X, Shu X, Wu X, et al. Dose-response relationship in non-pharmacological interventions for individuals with mild cognitive impairment: a systematic review and meta-analysis of randomised controlled trials. J Clin Nurs 2022 Dec;31(23-24):3390-3401. [doi: 10.1111/jocn.16240] [Medline: 35098610]
28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
30. A scoping review of non-pharmacological multi-modal interventions on multi-domain functions in older adults with mild cognitive impairment. Open Science Framework. URL: https://osf.io/y4hgk/ [accessed 2024-05-29]
31. Tricco AC, Lillie E, Zarin W, O'Brien K, Colquhoun H, Kastner M, et al. A scoping review on the conduct and reporting of scoping reviews. BMC Med Res Methodol 2016 Feb 09;16:15 [FREE Full text] [doi: 10.1186/s12874-016-0116-4] [Medline: 26857112]
32. Hagovska M, Nagyova I. The transfer of skills from cognitive and physical training to activities of daily living: a randomised controlled study. Eur J Ageing 2017 Jun;14(2):133-142 [FREE Full text] [doi: 10.1007/s10433-016-0395-y] [Medline: 28804398]
33. Hagovská M, Olekszyová Z. Relationships between balance control and cognitive functions, gait speed, and activities of daily living. Z Gerontol Geriatr 2016 Jul;49(5):379-385. [doi: 10.1007/s00391-015-0955-3] [Medline: 26458911]
34. Hagovská M, Olekszyová Z. Impact of the combination of cognitive and balance training on gait, fear and risk of falling and quality of life in seniors with mild cognitive impairment. Geriatr Gerontol Int 2016 Sep;16(9):1043-1050. [doi: 10.1111/ggi.12593] [Medline: 26338465]
35. Hagovská M, Takáč P, Dzvoník O. Effect of a combining cognitive and balanced training on the cognitive, postural and functional status of seniors with a mild cognitive deficit in a randomized, controlled trial. Eur J Phys Rehabil Med 2016 Feb;52(1):101-109 [FREE Full text] [Medline: 26325026]
36. Liao YY, Tseng HY, Lin YJ, Wang CJ, Hsu WC. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. Eur J Phys Rehabil Med 2020 Feb;56(1):47-57 [FREE Full text] [doi: 10.23736/S1973-9087.19.05899-4] [Medline: 31615196]
37. Liao YY, Chen IH, Lin YJ, Chen Y, Hsu WC. Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2019 Jul 16;11:162 [FREE Full text] [doi: 10.3389/fnagi.2019.00162] [Medline: 31379553]
38. Fiatarone Singh MA, Gates N, Saigal N, Wilson GC, Meiklejohn J, Brodaty H, et al. The Study of Mental and Resistance Training (SMART) study—resistance training and/or cognitive training in mild cognitive impairment: a randomized, double-blind, double-sham controlled trial. J Am Med Dir Assoc 2014 Dec;15(12):873-880. [doi: 10.1016/j.jamda.2014.09.010] [Medline: 25444575]
39. Fogarty JN, Murphy KJ, McFarlane B, Montero-Odasso M, Wells J, Troyer AK, et al. Taoist Tai Chi® and memory intervention for individuals with mild cognitive impairment. J Aging Phys Act 2016 Apr;24(2):169-180 [FREE Full text] [doi: 10.1123/japa.2014-0062] [Medline: 25838271]
40. Jeong MK, Park KW, Ryu JK, Kim GM, Jung HH, Park H. Multi-component intervention program on habitual physical activity parameters and cognitive function in patients with mild cognitive impairment: a randomized controlled trial. Int J Environ Res Public Health 2021 Jun 09;18(12):6240 [FREE Full text] [doi: 10.3390/ijerph18126240] [Medline: 34207701]
41. Kamegaya T, Araki Y, Kigure H, Long-Term-Care Prevention Team of Maebashi City, Yamaguchi H. Twelve-week physical and leisure activity programme improved cognitive function in community-dwelling elderly subjects: a randomized controlled trial. Psychogeriatrics 2014 Mar 16;14(1):47-54. [doi: 10.1111/psyg.12038] [Medline: 24528600]
42. Kim JH, Han JY, Park GC, Lee JS. Cognitive improvement effects of electroacupuncture combined with computer-based cognitive rehabilitation in patients with mild cognitive impairment: a randomized controlled trial. Brain Sci 2020 Dec 14;10(12):984 [FREE Full text] [doi: 10.3390/brainsci10120984] [Medline: 33327627]
43. Köbe T, Witte AV, Schnelle A, Lesemann A, Fabian S, Tesky VA, et al. Combined omega-3 fatty acids, aerobic exercise and cognitive stimulation prevents decline in gray matter volume of the frontal, parietal and cingulate cortex in patients with mild cognitive impairment. Neuroimage 2016 May 01;131:226-238. [doi: 10.1016/j.neuroimage.2015.09.050] [Medline: 26433119]
44. Lau CI, Liu MN, Cheng FY, Wang HC, Walsh V, Liao YY. Can transcranial direct current stimulation combined with interactive computerized cognitive training boost cognition and gait performance in older adults with mild cognitive impairment?: a randomized controlled trial. J Neuroeng Rehabil 2024 Feb 16;21(1):26 [FREE Full text] [doi: 10.1186/s12984-024-01313-0] [Medline: 38365761]
45. Liao YY, Liu MN, Wang HC, Walsh V, Lau CI. Combining transcranial direct current stimulation with tai chi to improve dual-task gait performance in older adults with mild cognitive impairment: a randomized controlled trial. Front Aging Neurosci 2021;13:766649 [FREE Full text] [doi: 10.3389/fnagi.2021.766649] [Medline: 34966268]

46. Phoemsapthawee J, Ammawat W, Prasertsri P, Sathalalai P, Leelayuwat N. Does Gotu kola supplementation improve cognitive function, inflammation, and oxidative stress more than multicomponent exercise alone?: a randomized controlled study. J Exerc Rehabil 2022 Oct;18(5):330-342 [FREE Full text] [doi: 10.12965/jer.2244388.194] [Medline: 36420467]
47. Senczyszyn A, Szcześniak D, Wieczorek T, Maciaszek J, Małecka M, Bogudzińska B, et al. Improvement of working memory in older adults with mild cognitive impairment after repetitive transcranial magnetic stimulation - a randomized controlled pilot study. Front Psychiatry 2023;14:1196478 [FREE Full text] [doi: 10.3389/fpsyt.2023.1196478] [Medline: 38111617]
48. Shimizu N, Umemura T, Matsunaga M, Hirai T. Effects of movement music therapy with a percussion instrument on physical and frontal lobe function in older adults with mild cognitive impairment: a randomized controlled trial. Aging Ment Health 2018 Dec;22(12):1614-1626 [FREE Full text] [doi: 10.1080/13607863.2017.1379048] [Medline: 28937272]
49. Yang QH, Lyu X, Lin QR, Wang ZW, Tang L, Zhao Y, et al. Effects of a multicomponent intervention to slow mild cognitive impairment progression: a randomized controlled trial. Int J Nurs Stud 2022 Jan;125:104110. [doi: 10.1016/j.ijnurstu.2021.104110] [Medline: 34736073]
50. Bray NW, Pieruccini-Faria F, Witt ST, Bartha R, Doherty TJ, Nagamatsu LS, et al. Combining exercise with cognitive training and vitamin D 3 to improve functional brain connectivity (FBC) in older adults with mild cognitive impairment (MCI). Results from the SYNERGIC trial. Geroscience 2023 Jun;45(3):1967-1985 [FREE Full text] [doi: 10.1007/s11357-023-00805-6] [Medline: 37162700]
51. Fairchild JK, Myers J, Louras P, Jo B, McNerney MW, Hallmayer J, et al. Multimodal exercise and cognitive training program improves cognitive function in amnestic mild cognitive impairment. Am J Geriatr Psychiatry 2024 Apr;32(4):463-474. [doi: 10.1016/j.jagp.2023.12.002] [Medline: 38220592]
52. Griffiths J, Thaikruea L, Wongpakaran N, Munkhetvit P, Kittisares A, Varnado P. Effects of combined physical movement activity and multifaceted cognitive training in older people with mild neurocognitive disorder in a rural community: a randomized control trial. Dement Geriatr Cogn Disord 2020;49(2):194-201. [doi: 10.1159/000507922] [Medline: 32535601]
53. Martin DM, Mohan A, Alonzo A, Gates N, Gbadeyan O, Meinzer M, et al. A pilot double-blind randomized controlled trial of cognitive training combined with transcranial direct current stimulation for amnestic mild cognitive impairment. J Alzheimers Dis 2019 Aug 12;71(2):503-512. [doi: 10.3233/JAD-190306] [Medline: 31424410]
54. Montero-Odasso M, Zou G, Speechley M, Almeida QJ, Liu-Ambrose T, Middleton LE, Canadian GaitCognition Network. Effects of exercise alone or combined with cognitive training and vitamin D supplementation to improve cognition in adults with mild cognitive impairment: a randomized clinical trial. JAMA Netw Open 2023 Jul 03;6(7):e2324465 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.24465] [Medline: 37471089]
55. Bae S, Lee S, Lee S, Jung S, Makino K, Harada K, et al. The effect of a multicomponent intervention to promote community activity on cognitive function in older adults with mild cognitive impairment: a randomized controlled trial. Complement Ther Med 2019 Feb;42:164-169. [doi: 10.1016/j.ctim.2018.11.011] [Medline: 30670238]
56. Blumenthal JA, Smith PJ, Mabe S, Hinderliter A, Lin P, Liao L, et al. Lifestyle and neurocognition in older adults with cognitive impairments: a randomized trial. Neurology 2019 Jan 15;92(3):e212-e223 [FREE Full text] [doi: 10.1212/WNL.0000000000006784] [Medline: 30568005]
57. Callisaya ML, Jayakody O, Vaidya A, Srikanth V, Farrow M, Delbaere K. A novel cognitive-motor exercise program delivered via a tablet to improve mobility in older people with cognitive impairment - StandingTall cognition and mobility. Exp Gerontol 2021 Sep;152:111434. [doi: 10.1016/j.exger.2021.111434] [Medline: 34098009]
58. Chobe S, Patra SK, Chobe M, Metri K. Efficacy of integrated yoga and ayurveda rasayana on cognitive functions in elderly with mild cognitive impairment: non-randomized three-arm clinical trial. J Ayurveda Integr Med 2022;13(1):100373 [FREE Full text] [doi: 10.1016/j.jaim.2020.11.003] [Medline: 33349568]
59. Gill DP, Gregory MA, Zou G, Liu-Ambrose T, Shigematsu R, Hachinski V, et al. The healthy mind, healthy mobility trial: a novel exercise program for older adults. Med Sci Sports Exerc 2016 Feb;48(2):297-306. [doi: 10.1249/MSS.0000000000000758] [Medline: 26285025]
60. Gonzalez PC, Fong KN, Brown T. Transcranial direct current stimulation as an adjunct to cognitive training for older adults with mild cognitive impairment: a randomized controlled trial. Ann Phys Rehabil Med 2021 Sep;64(5):101536 [FREE Full text] [doi: 10.1016/j.rehab.2021.101536] [Medline: 33957292]
61. Jesus AG, de Lima MP, Vilar M, Pachana NA. Feasibility, acceptability, and preliminary efficacy of a cognitive-emotional intervention program in homebound older adults: a pilot randomized controlled trial. Clin Gerontol 2023;46(3):359-375. [doi: 10.1080/07317115.2022.2105670] [Medline: 35916398]
62. Jones KT, Ostrand AE, Gazzaley A, Zanto TP. Enhancing cognitive control in amnestic mild cognitive impairment via at-home non-invasive neuromodulation in a randomized trial. Sci Rep 2023 May 08;13(1):7435 [FREE Full text] [doi: 10.1038/s41598-023-34582-1] [Medline: 37156876]
63. Kim JH, Park JH. Does cognitive-physical dual-task training have better clinical outcomes than cognitive single-task training does? A single-blind, randomized controlled trial. Healthcare (Basel) 2023 May 25;11(11):1544 [FREE Full text] [doi: 10.3390/healthcare11111544] [Medline: 37297684]

64. Straubmeier M, Behrndt EM, Seidl H, Özbe D, Luttenberger K, Graessel E. Non-pharmacological treatment in people with cognitive impairment. Dtsch Arztebl Int 2017 Dec 01;114(48):815-821 [FREE Full text] [doi: 10.3238/arztebl.2017.0815] [Medline: 29249224]
65. Xu Z, Zhang D, Lee AT, Sit RW, Wong C, Lee EK, et al. A pilot feasibility randomized controlled trial on combining mind-body physical exercise, cognitive training, and nurse-led risk factor modification to reduce cognitive decline among older adults with mild cognitive impairment in primary care. PeerJ 2020;8:e9845 [FREE Full text] [doi: 10.7717/peerj.9845] [Medline: 33194354]
66. Li B, Tang H, He G, Jin Z, He Y, Huang P, et al. Tai Chi enhances cognitive training effects on delaying cognitive decline in mild cognitive impairment. Alzheimers Dement 2023 Jan 15;19(1):136-149. [doi: 10.1002/alz.12658] [Medline: 35290704]
67. Xu Y, Zhu J, Liu H, Qiu Z, Wu M, Liu J, et al. Effects of Tai Chi combined with tDCS on cognitive function in patients with MCI: a randomized controlled trial. Front Public Health 2023;11:1199246 [FREE Full text] [doi: 10.3389/fpubh.2023.1199246] [Medline: 37608981]
68. Stuerz K, Hartmann S, Holzner B, Bichler CS, Niedermeier M, Kopp M, et al. Effects of a two-step cognitive and relaxation training program in care home residents with mild cognitive impairment. Int J Environ Res Public Health 2022 Jul 07;19(14):8316 [FREE Full text] [doi: 10.3390/ijerph19148316] [Medline: 35886176]
69. Jesus AG, Lima MP, Vilar M, Pachana NA. HEPPI: a randomized controlled trial of the efficacy of a cognitive-emotional intervention for homebound older adults. Aging Ment Health 2024;28(3):491-501. [doi: 10.1080/13607863.2023.2260760] [Medline: 37747057]
70. Rezola-Pardo C, Arrieta H, Gil SM, Zarrazquin I, Yanguas JJ, López MA, et al. Comparison between multicomponent and simultaneous dual-task exercise interventions in long-term nursing home residents: the Ageing-ONDUAL-TASK randomized controlled study. Age Ageing 2019 Nov 01;48(6):817-823. [doi: 10.1093/ageing/afz105] [Medline: 31595289]
71. Shimada H, Makizako H, Doi T, Park H, Tsutsumimoto K, Verghese J, et al. Effects of combined physical and cognitive exercises on cognition and mobility in patients with mild cognitive impairment: a randomized clinical trial. J Am Med Dir Assoc 2018 Jul;19(7):584-591. [doi: 10.1016/j.jamda.2017.09.019] [Medline: 29153754]
72. Park H, Park JH, Na HR, Hiroyuki S, Kim GM, Jung MK, et al. Combined intervention of physical activity, aerobic exercise, and cognitive exercise intervention to prevent cognitive decline for patients with mild cognitive impairment: a randomized controlled clinical study. J Clin Med 2019 Jun 28;8(7):940 [FREE Full text] [doi: 10.3390/jcm8070940] [Medline: 31261796]
73. Tao M, Liu H, Cheng J, Yu C, Zhao L. Motor-cognitive interventions may effectively improve cognitive function in older adults with mild cognitive impairment: a randomized controlled trial. Behav Sci (Basel) 2023 Sep 04;13(9):737 [FREE Full text] [doi: 10.3390/bs13090737] [Medline: 37754015]
74. Lam LC, Chan WC, Leung T, Fung AW, Leung EM. Would older adults with mild cognitive impairment adhere to and benefit from a structured lifestyle activity intervention to enhance cognition?: a cluster randomized controlled trial. PLoS One 2015;10(3):e0118173 [FREE Full text] [doi: 10.1371/journal.pone.0118173] [Medline: 25826620]
75. Lee S, Harada K, Bae S, Harada K, Makino K, Anan Y, et al. A non-pharmacological multidomain intervention of dual-task exercise and social activity affects the cognitive function in community-dwelling older adults with mild to moderate cognitive decline: a randomized controlled trial. Front Aging Neurosci 2023;15:1005410 [FREE Full text] [doi: 10.3389/fnagi.2023.1005410] [Medline: 36993908]
76. Lin YC, Wang CJ, Chang YF, Wang JJ. Effects of the biopsychosocial functional activity program on cognitive function for community older adults with mild cognitive impairment: a cluster-randomized controlled trial. Nurs Health Sci 2020 Dec;22(4):1065-1075. [doi: 10.1111/nhs.12772] [Medline: 32865857]
77. Train the Brain Consortium. Randomized trial on the effects of a combined physical/cognitive training in aged MCI subjects: the Train the Brain study. Sci Rep 2017 Jan 03;7:39471 [FREE Full text] [doi: 10.1038/srep39471] [Medline: 28045051]
78. Liu X, Ma Z, Zhu X, Zheng Z, Li J, Fu J, et al. Cognitive benefit of a multidomain intervention for older adults at risk of cognitive decline: a cluster-randomized controlled trial. Am J Geriatr Psychiatry 2023 Mar;31(3):197-209. [doi: 10.1016/j.jagp.2022.10.006] [Medline: 36414488]
79. Sungkarat S, Boripuntakul S, Kumfu S, Lord SR, Chattipakorn N. Tai Chi improves cognition and plasma BDNF in older adults with mild cognitive impairment: a randomized controlled trial. Neurorehabil Neural Repair 2018 Feb;32(2):142-149. [doi: 10.1177/1545968317753682] [Medline: 29353543]
80. Barnes LL, Dhana K, Liu X, Carey VJ, Ventrelle J, Johnson K, et al. Trial of the MIND diet for prevention of cognitive decline in older persons. N Engl J Med 2023 Aug 17;389(7):602-611 [FREE Full text] [doi: 10.1056/NEJMoa2302368] [Medline: 37466280]
81. Andrews V, Zammit G, O'Leary F. Dietary pattern, food, and nutritional supplement effects on cognitive outcomes in mild cognitive impairment: a systematic review of previous reviews. Nutr Rev 2023 Oct 10;81(11):1462-1489 [FREE Full text] [doi: 10.1093/nutrit/nuad013] [Medline: 37027832]
82. McGrattan AM, McEvoy CT, McGuinness B, McKinley MC, Woodside JV. Effect of dietary interventions in mild cognitive impairment: a systematic review. Br J Nutr 2018 Dec;120(12):1388-1405 [FREE Full text] [doi: 10.1017/S0007114518002945] [Medline: 30409231]
83. Vlachos GS, Scarmeas N. Dietary interventions in mild cognitive impairment and dementia. Dialogues Clin Neurosci 2019 Mar;21(1):69-82 [FREE Full text] [doi: 10.31887/DCNS.2019.21.1/nscarmeas] [Medline: 31607782]

84. Kuiper JS, Zuidersma M, Zuidema SU, Burgerhof JG, Stolk RP, Oude Voshaar RC, et al. Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies. Int J Epidemiol 2016 Aug 06;45(4):1169-1206. [doi: 10.1093/ije/dyw089] [Medline: 27272181]
85. Wang HX, Karp A, Winblad B, Fratiglioni L. Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project. Am J Epidemiol 2002 Jun 15;155(12):1081-1087. [doi: 10.1093/aje/155.12.1081] [Medline: 12048221]
86. Schneider N, Yvon C. A review of multidomain interventions to support healthy cognitive ageing. J Nutr Health Aging 2013 Mar;17(3):252-257 [FREE Full text] [doi: 10.1007/s12603-012-0402-8] [Medline: 23459978]
87. Hill NT, Mowszowski L, Naismith SL, Chadwick VL, Valenzuela M, Lampit A. Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. Am J Psychiatry 2017 Apr 01;174(4):329-340. [doi: 10.1176/appi.ajp.2016.16030360] [Medline: 27838936]
88. Wu J, Ma Y, Ren Z. Rehabilitative effects of virtual reality technology for mild cognitive impairment: a systematic review with meta-analysis. Front Psychol 2020 Sep 25;11:1811 [FREE Full text] [doi: 10.3389/fpsyg.2020.01811] [Medline: 33101098]
89. Thapa N, Park HJ, Yang JG, Son H, Jang M, Lee J, et al. The effect of a virtual reality-based intervention program on cognition in older adults with mild cognitive impairment: a randomized control trial. J Clin Med 2020 Apr 29;9(5):1283 [FREE Full text] [doi: 10.3390/jcm9051283] [Medline: 32365533]
90. D'Cunha NM, Nguyen D, Naumovski N, McKune AJ, Kellett J, Georgousopoulou EN, et al. A mini-review of virtual reality-based interventions to promote well-being for people living with dementia and mild cognitive impairment. Gerontology 2019;65(4):430-440. [doi: 10.1159/000500040] [Medline: 31108489]
91. Rosenberg A, Mangialasche F, Ngandu T, Solomon A, Kivipelto M. Multidomain interventions to prevent cognitive impairment, Alzheimer's disease, and dementia: from FINGER to world-wide FINGERS. J Prev Alzheimers Dis 2020;7(1):29-36 [FREE Full text] [doi: 10.14283/jpad.2019.41] [Medline: 32010923]
92. Hugo J, Ganguli M. Dementia and cognitive impairment: epidemiology, diagnosis, and treatment. Clin Geriatr Med 2014 Aug;30(3):421-442 [FREE Full text] [doi: 10.1016/j.cger.2014.04.001] [Medline: 25037289]
93. Lu Y, An Y, Guo J, Zhang X, Wang H, Rong H, et al. Dietary intake of nutrients and lifestyle affect the risk of mild cognitive impairment in the Chinese elderly population: a cross-sectional study. Front Behav Neurosci 2016;10:229 [FREE Full text] [doi: 10.3389/fnbeh.2016.00229] [Medline: 27965552]
94. Smith PJ, Blumenthal JA. Dietary factors and cognitive decline. J Prev Alzheimers Dis 2016 Mar;3(1):53-64 [FREE Full text] [doi: 10.14283/jpad.2015.71] [Medline: 26900574]

Abbreviations

CCT: computerized cognitive training
CT: cognitive training
DLPFC: dorsolateral prefrontal cortex
MCI: mild cognitive impairment
NI: nutrition intervention
PI: psychosocial intervention
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PT: physical training
SA: social activities
tACS: transcranial alternating current stimulation
tDCS: transcranial direct current stimulation
VR: virtual reality

Edited by H LaMonica; submitted 19.12.24; peer-reviewed by F Shojaei, D Tao, S Moolsart; comments to author 16.02.25; revised version received 07.04.25; accepted 21.04.25; published 12.05.25.

Please cite as:
Chan RCF, Zhou JHS, Cao Y, Lo K, Ng PHF, Shum DHK, Wong AYL
Nonpharmacological Multimodal Interventions for Cognitive Functions in Older Adults With Mild Cognitive Impairment: Scoping Review
JMIR Aging 2025;8:e70291
URL: https://aging.jmir.org/2025/1/e70291
doi:10.2196/70291
PMID:

Review

Advancing Remote Monitoring for Patients With Alzheimer Disease and Related Dementias: Systematic Review

Mohmmad Arif Shaik[1], MS; Fahim Islam Anik[2], MS; Md. Mehedi Hasan[1], MS; Sumit Chakravarty[1], PhD; Mary Dioise Ramos[3], RN, PhD; Mohammad Ashiqur Rahman[4], PhD; Sheikh Iqbal Ahamed[5], PhD; Nazmus Sakib[1], PhD

[1]Department of Electrical and Computer Engineering, Kennesaw State University, Kennesaw, GA, United States

[2]Department of Mechanical Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh

[3]Louisiana State University Health Sciences Center (LSUHSC) New Orleans, School of Nursing, Louisiana State University, New Orleans, United States

[4]Department of Electrical and Computer Engineering, Florida International University, Miami, FL, United States

[5]Department of Computer Science, Marquette University, Milwaukee, WI, United States

Corresponding Author:
Mohmmad Arif Shaik, MS
Department of Electrical and Computer Engineering
Kennesaw State University
Kennesaw, GA
United States
Phone: 1 4147975981
Email: mshaik14@students.kennesaw.edu

Abstract

Background: Using remote monitoring technology in the context of Alzheimer disease (AD) care presents exciting new opportunities to lessen caregiver stress and improve patient care quality. The application of wearables, environmental sensors, and smart home systems designed specifically for patients with AD represents a promising interdisciplinary approach that integrates advanced technology with health care to enhance patient safety, monitor health parameters in real time, and provide comprehensive support to caregivers.

Objective: The objectives of this study included evaluating the effectiveness of various remote sensing technologies in enhancing patient outcomes and identifying strategies to alleviate the burden on health care professionals and caregivers. Critical elements such as regulatory compliance, user-centered design, privacy and security considerations, and the overall efficacy of relevant technologies were comprehensively examined. Ultimately, this study aimed to propose a comprehensive remote monitoring framework tailored to the needs of patients with AD and related dementias.

Methods: Guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework, we conducted a systematic review on remote monitoring for patients with AD and related dementias. Our search spanned 4 major electronic databases—Google Scholar, PubMed, IEEE Xplore, and DBLP on February 20, 2024, with an updated search on May 18, 2024.

Results: A total of 31 publications met the inclusion criteria, highlighting 4 key research areas: existing remote monitoring technologies, balancing practicality and empathy, security and privacy in monitoring, and technology design for AD care. The studies revealed a strong focus on various remote monitoring methods for capturing behavioral, physiological, and environmental data yet showed a gap in evaluating these methods for patient and caregiver needs, privacy, and usability. The findings also indicated that many studies lacked robust reference standards and did not consistently apply critical appraisal criteria, underlining the need for comprehensive frameworks that better integrate these essential considerations.

Conclusions: This comprehensive literature review of remote monitoring technologies for patients with AD provides an understanding of remote monitoring technologies, trends, and gaps in the current research and the significance of novel strategies for remote monitoring to enhance patient outcomes and reduce the burden among health professionals and caregivers. The proposed remote monitoring framework aims to inspire the development of new interdisciplinary research models that advance care for patients with AD.

(JMIR Aging 2025;8:e69175) doi:10.2196/69175

KEYWORDS

dementia; Alzheimer disease; remote monitoring; Alzheimer; caregiver; fall detection; artificial intelligence

Introduction

Background

Alzheimer disease (AD) is the most generic form of dementia, and significantly affects the brain cells, declines cognitive abilities, and makes it difficult to perform daily tasks [1]. The main effects of the disease are loss of memory, inability to use problem-solving and logical thinking skills, anxiety, depression, and confusion [2]. AD has remained the leading cause of death among people aged >60 years. The main cause of AD is the accumulation of abnormal proteins such as β-amyloid and microtubule-associated tau protein in the brain; the aggregation of these proteins causes synaptic disintegration and neural loss, which characterize AD [3,4]. Disturbingly, it is estimated that the number of AD cases could grow to 13.8 million by 2060. The pervasive nature of this disease underscores the pressing need for continued research, intervention strategies, and support systems to address the growing challenges posed by AD in an aging population [5,6].

Caring for individuals with AD is an enduring and demanding journey, often placing a significant burden on family members who become primary caregivers or on those who opt to hire professional caregivers [7]. The responsibilities encompass a spectrum of care, requiring a deep commitment to providing the essential support needed for patients with AD to navigate their daily lives. Caregivers play a crucial role in assisting individuals with AD in tasks integral to their routine while also implementing lifestyle adjustments aimed at alleviating stress, confusion, anxiety, and agitation. The caregiving journey involves fostering an environment that promotes mental, physical, and social well-being. This includes implementing tailored strategies to keep patients with AD mentally engaged, encouraging physical activities that align with their abilities, and facilitating social interactions to prevent isolation. However, continuous monitoring and provision of care for patients with AD causes a substantial burden on both the family members and caregivers [8,9]. This responsibility often causes a lot of stress and depression among those providing care. Family members and caregivers find themselves dealing with the progressive loss of individuals with AD. In addition to that, they must navigate between caregiver responsibilities and sustaining a work-life balance. Disputes between family members, financial burdens, and escalating medical expenses extend the challenges that those undertaking caregiver roles face [10]. While existing research has identified various scales to measure the burden on family members, there is a notable gap in the literature concerning effective strategies to alleviate this burden. Manzini and do Vale [11] and Liao et al [12] address this critical gap by exploring remote monitoring methods facilitated by ITs. By leveraging innovative approaches, the research seeks to enhance the quality of care provided to patients with AD and substantially reduce the burden experienced by family members and caregivers. Through the implementation of such technological solutions, the aforementioned studies [11,12] aspired to contribute valuable insights into improving the overall well-being of both patients with AD and those entrusted with their care.

Despite the growing body of research on remote monitoring for AD and related dementias (ADRD), significant gaps remain in the literature, particularly regarding interdisciplinary integration. Previous reviews have primarily focused on individual aspects of remote monitoring, such as wearable devices [13-16], smart home sensors [17-19], or privacy and security considerations [20,21], often in isolation. However, the effectiveness of these technologies depends on a holistic approach that integrates expertise from health care, computer science, cybersecurity, and user-centered design. The lack of interdisciplinary research has led to fragmented solutions that may not fully address the practical needs of patients and caregivers. This review sought to bridge these gaps by offering a comprehensive assessment of remote monitoring technologies while emphasizing their usability, security, and ethical implications. Therefore, we formulated four research questions (RQs) to explore existing remote monitoring technologies and their applications, assess their needs and advantages, examine privacy concerns related to remote monitoring, and investigate technology adoption among older adults:

- RQ 1: What are the primary remote monitoring technologies to capture different behavioral, physiological, and environmental data?
- RQ 2: To what extent do these existing remote monitoring technologies achieve an optimal balance of efficiency, practicality, and empathy in addressing the needs of both patients and caregivers?
- RQ 3: What are the critical security and privacy considerations and concerns involved in these remote monitoring technologies?
- RQ 4: What are the critical design considerations and technology components for a comprehensive remote monitoring framework specifically tailored to patients with AD?

While this review focused on remote monitoring for individuals with AD, it is important to acknowledge that many of the challenges, technologies, and solutions discussed are also relevant to other dementia subtypes. Conditions such as vascular dementia [22,23], Lewy body dementia [24], and frontotemporal dementia [25] share overlapping symptoms [26,27], including cognitive decline, behavior changes, and mobility issues, which can similarly benefit from remote monitoring interventions. However, our decision to center this review on AD was guided by its high prevalence and the significant body of research available on remote monitoring solutions tailored specifically to ADRD. The insights presented in this paper may be broadly applicable to dementia care, although future research should further explore technology adaptations for different dementia subtypes.

The remainder of this paper is divided into 3 key sections. Section 2 outlines the methodology, comprising 3 subsections that detail the scoping criteria, literature search strategy, and

data analysis procedures. Section 3 addresses the RQs in depth. Finally, section 4 concludes the literature review, summarizing the key findings and their implications.

Objectives

Our objective with this paper was to critically assess the effectiveness of various remote sensing technologies in enhancing patient outcomes, particularly for individuals with ADRD, while also identifying strategies to reduce the burden on health care professionals (HCPs) and caregivers. To perform this assessment, the research comprehensively examined key factors essential for the successful implementation and adoption of these technologies. These factors include patient and caregiver compliance, ensuring that the technologies are used consistently and effectively; user-centered design, which focuses on creating accessible and intuitive systems tailored to the specific needs of end users; privacy and security concerns, addressing the safeguarding of sensitive patient data and ensuring adherence to ethical and regulatory standards; and the overall efficacy of these technologies, evaluating their accuracy, reliability, and impact on patient care.

Methods

Overview

In the landscape of AD care, this review undertook a comprehensive examination of the complex dynamics introduced by remote monitoring technologies. We focused on different remote monitoring technologies, types of algorithms used to process the data, issues and concerns related to security and privacy, caregiver burdens, and specific monitoring frameworks for patients with AD. We conducted an extensive search to identify relevant research papers across multiple disciplines from 2011 to 2024. Our search encompassed a diverse array of digital libraries, including Google Scholar, IEEE Xplore, PubMed, and DBLP on February 20, 2024, with an updated search on May 18, 2024, to ensure coverage across multidisciplinary domains, including health care, engineering, and computer science. These databases were selected based on their relevance to remote monitoring technologies, security and privacy considerations, and AD care. While Scopus, Embase, and Cochrane are widely recognized for systematic reviews in clinical and biomedical research, our focus on technological, interdisciplinary, and engineering aspects made IEEE Xplore and DBLP particularly valuable sources. In addition, PubMed was included to ensure coverage of relevant medical and health care studies. Google Scholar was used to capture a broader range of studies, including conference proceedings and gray literature, which are crucial for emerging technologies. On Google Scholar, the search string used was *"Remote monitoring technologies" AND "Wearable sensors" AND Alzheimer AND "Privacy concerns,"* ensuring a broad examination of related literature. PubMed and IEEE Xplore were queried using the search string *(Remote monitoring technologies) AND Alzheimer AND ("Privacy concerns")* to capture articles with a focus on technical and medical perspectives. In addition, DBLP was queried using the search string *Alzheimer AND "Remote monitoring"* to include publications emphasizing computer science and engineering insights. In addition, we cross-referenced the citations of each article to uncover additional relevant articles or government reports aligning with our inclusion criteria. Notably, we imposed some constraints on the period during which the papers were published, and only English-language papers were considered. In total, we identified 31 articles that fulfilled our inclusion criteria.

Our search strategy included keywords such as "Remote monitoring technologies" AND "Wearable sensors" AND "Alzheimer" AND "Privacy concerns" to ensure the inclusion of studies addressing security and data protection—critical challenges in adopting remote monitoring solutions. However, we acknowledge that this keyword choice may have limited the retrieval of studies focusing on broader aspects, such as clinical effectiveness, usability, and caregiver perspectives. While our selection process incorporated additional screening to identify studies covering a wider range of topics, future research could benefit from refining the search terms to ensure a more balanced representation of different remote monitoring challenges.

Study Identification and Selection

A total of 79 papers were initially identified based on their titles and abstracts that were deemed to be potentially relevant to the research topic. Of the 61 records assessed for eligibility, 30 (49%) were excluded for various reasons: 7 (23%) were books, 6 (20%) were partial articles, and 17 (57%) were out-of-context studies. Finally, 31 papers were identified to extract information and insights related to the RQs. These papers were thoroughly analyzed to identify common themes, methodologies, findings, and gaps in the existing literature. This analysis provided a deeper understanding of the current research and valuable insights to guide this review. This systematic approach to literature review and analysis enhanced the validity and reliability of the research, contributing to its overall credibility.

In the following sections, the findings from the papers are presented and discussed. The results are organized based on the information identified in the literature review with respect to the RQs.

Results and Discussion

Primary Remote Monitoring Technologies (RQ 1)

Overview

Remote monitoring technologies using wearable and portable device–based sensors have revolutionized health care by enabling continuous and real-time tracking of various physiological and environmental parameters. These sensors, embedded in devices such as smartwatches, fitness bands, and portable health monitors, can measure heart rate, blood pressure, glucose levels, temperature, and physical activity [28,29]. They offer significant advantages in managing chronic conditions, detecting early signs of health deterioration, and providing personalized care. Wearable sensors are particularly beneficial in monitoring older patients and individuals with chronic diseases, allowing for prompt interventions and reducing hospital visits. Furthermore, the integration of these devices with mobile health (mHealth) apps and cloud-based platforms facilitates seamless data collection, analysis, and sharing with

health care providers, enhancing patient engagement and improving health outcomes. The subsequent subsections provide a detailed description of the various aspects of remote monitoring technologies. To ensure a systematic approach, a thorough review of the literature was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework [30,31] (Multimedia Appendix 1), as shown in Figure 1. Figure 2 illustrates the integration and synergy of different remote monitoring technologies based on relevant parameters. These are further elaborated on in the following subsections and in Multimedia Appendices 2-5 [14,32-52].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the study selection in the literature review.

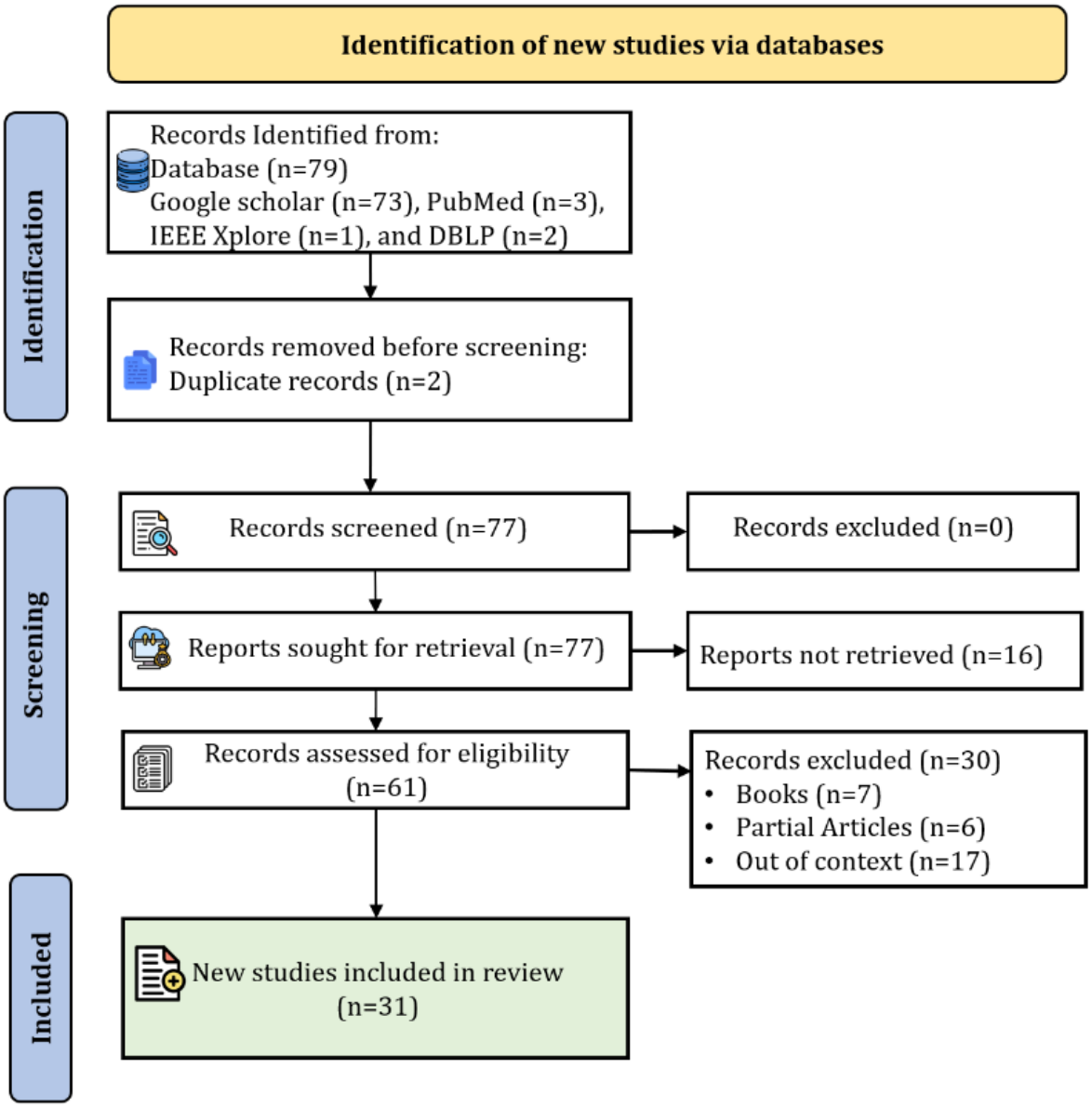
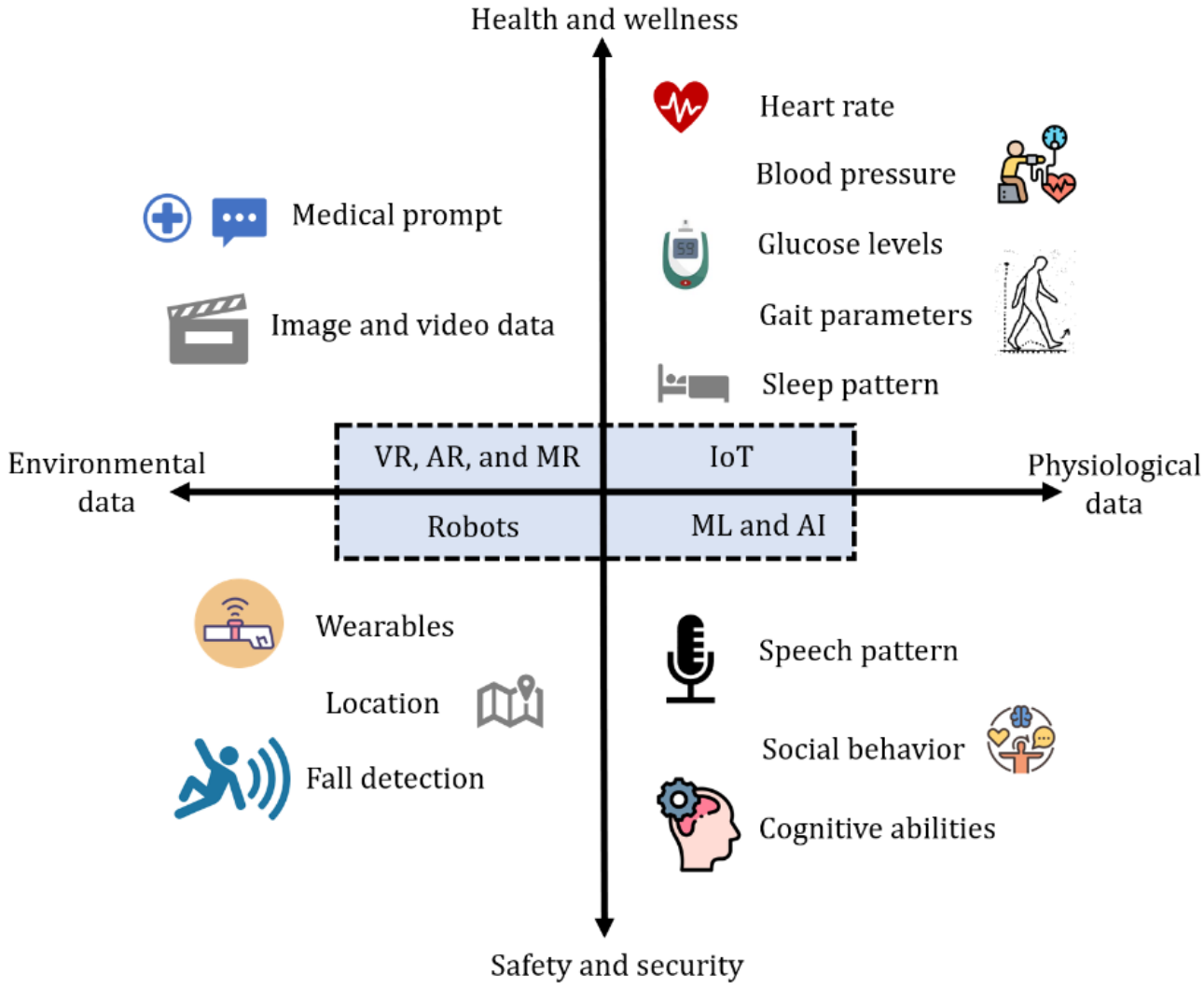


Figure 2. Integration and synergy of different remote monitoring technologies. AI: artificial intelligence; AR: augmented reality; IoT: Internet of Things; ML: machine learning; MR: mixed reality; VR: virtual reality.



Wearables

Wearable technologies refer to devices worn on the body either as an accessory or as clothing. Devices embedded with an accelerometer and gyroscope, such as wrist or ankle devices, are used for analyzing gait parameters such as speed, variability, and stride length [32]. There are also wearable devices that capture physical activity data, including steps taken, intensity, and duration, aiding in tracking physical activity, which helps the patient with AD be physically active and slow down the progression of the disease [32]. In addition, wrist devices and smart clothes are equipped with sensors, whereas various wearables, including wristbands, smartwatches, and chest straps, collect different data such as physiological signs, posture, gait, movement, and sleep data, and enable the continuous monitoring of blood pressure. Devices such as the Basis Health Tracker, Misfit Shine, Fitbit Flex, Withings Pulse O2, and Actiwatch Spectrum actigraph are used for sleep detection and monitoring daily activities [53-57]. These devices enable continuous monitoring of vulnerable patients, offering medical specialists a comprehensive view of the patients' health for more accurate diagnosis and treatment [33,34]. Moreover, dream headbands are used to record and store physiological data in real time during sleep, facilitating the monitoring of daily sleep patterns and enabling patients and caregivers to access their reports freely [35]. Further details are provided in Multimedia Appendix 2 [32-35,37-45] and Multimedia Appendix 3 [32-34,36-38,42,45-47].

Environment and Home Sensors

Remote monitoring through advanced environment and sensor technology provides a promising approach to enhancing care, monitoring daily activities, and detecting early signs of cognitive decline. In the context of home-based remote monitoring, various sensors are used to capture and analyze vital health and activity data. Infrared sensors strategically placed throughout the home facilitate gait analysis by detecting individual movements and providing insights into gait characteristics such as walking speed, rhythm, and variability [32]. Wireless mattress sensors and ambient sensors are used to monitor sleep time, movements, and patterns, aiding in the early detection of mild cognitive impairment (MCI) and AD. In addition, night monitoring systems equipped with motion detectors are used to monitor sleep patterns and alert caregivers to unusual nighttime behavior [32]. In the case of continuous physical activity monitoring, infrared sensors are used to collect internal

movement data, which allows for the assessment of activity level without the need for active participation of the individual [32,36]. A multisensory approach, including sound sensors, video cameras, radio-frequency identification readers, pressure sensors, light sensors, proximity sensors, temperature sensors, humidity sensors, and location trackers, has been applied to many activities and environmental conditions in the home [33]. This rich dataset facilitates the assessment of disease progression and provides valuable information about changes in daily activities and environmental factors. Behavior patterns are tracked using sensors installed in various areas of the home, including the kitchen, bedroom, and bathroom, along with fall detection sensors. These sensors help monitor mobility, stability, frequency, and time spent in different areas, contributing to a deeper understanding of daily routines and potential risks [32]. Microsoft Kinect cameras are specifically used for gait monitoring to detect and assess falls among older individuals, with the collected video data aiding in fall prevention efforts [37].

Finally, motion sensor–activated light paths and bed pressure sensors are used for fall prevention and sleep problem detection, respectively. These sensors provide an additional layer of safety and monitoring, enhancing the overall effectiveness of the remote monitoring system [34].

Personal Devices

Personal devices, specifically, smartphones equipped with advanced sensors such as accelerometers, gyroscopes, and magnetometers, play a pivotal role in gait analysis and physical activity monitoring. Apps installed on these smartphones analyze the sensor data to determine various gait characteristics, including stride length, walking speed, and rhythm, providing valuable insights into mobility patterns [32]. Moreover, waist-worn accelerometers facilitate continuous monitoring of physical activity parameters such as movements and acceleration, aiding individuals in maintaining optimal levels of physical activity and energy use [32]. In addition to physical activity monitoring, smartphones serve as effective tools for assessing social behavior. By monitoring calls, SMS text messages, and internet-browsing activities, these devices offer insights into an individual's social activity levels and patterns. This information is crucial for health care providers and caregivers to understand the social interactions of individuals, identify potential signs of isolation, and implement interventions to promote social engagement and well-being [37].

Machine Learning and Artificial Intelligence

Machine learning and artificial intelligence (AI) play a pivotal role in health care by leveraging diverse data sources to enhance diagnostic capabilities, predict disease progression, and support clinical decision-making. Driving patterns collected through GPSs are analyzed to detect early signs of preclinical AD. These patterns are compared with disease progression to identify behavior changes linked to AD [32]. Activity data from sensor-embedded devices such as wearables are used for data mining and predictive modeling, aiding in identifying potential health issues and informing health care decisions [46]. Sleep patterns collected from these devices offer valuable insights into an individual's health status, providing clinicians with early warnings of health deteriorations or relapses and enabling timely interventions to improve patient outcomes [47].

Speech pattern analysis using data from health records, including calls and SMS text messages, offers diagnostic capabilities for neurodegenerative diseases such as Parkinson disease and AD, enabling early detection and intervention [38]. In addition, AI-powered image and video processing techniques assist in diagnosing various conditions, such as cancer from x-ray and computed tomography scans and degenerative diseases from fundus photography, enhancing diagnostic accuracy and treatment planning [38]. Furthermore, data collected from wrist-worn actigraphy devices such as the ActiGraph wGT3X are analyzed using a variety of machine learning algorithms, including logistic regression, random forest, gradient-boosting machine, and support vector machine. These algorithms help predict the risks and progression of disease, providing valuable insights for personalized health care management and intervention strategies [45].

Robots

Robots such as Nao, Pepper, Paro, Stevie, Zoro, Mini, and Tangy are emerging as valuable tools in health care settings, particularly for supporting the needs of older adults. These robotic companions engage older individuals through interactive activities such as conversations, quizzes, tongue twisters, and arithmetic calculations, demonstrating promising results in enhancing social engagement and alleviating feelings of isolation [46]. AI-powered robots equipped with advanced sensors and algorithms can access environmental data and user-specific characteristics to deliver personalized medical prompts and emotional support. These robots play a crucial role in physical rehabilitation, reducing feelings of isolation and promoting overall well-being by providing companionship and interactive engagement [38,42]. Furthermore, specialized robots such as Nao, Pepper, and Paro are designed to cater to the unique needs of patients with dementia with cognitive impairment. These robots offer a comprehensive range of services, including behavioral monitoring, physical exercise tutoring, recreational activities, stress management, and medication reminders. By providing multifaceted support, these robotic companions significantly enhance the quality of life of older adults and contribute to improved health care outcomes [34]. Robots named Stevie and Zoro are specially tailored to support caregivers and older adults with memory loss. These robots offer care support, entertainment options, cognitive engagement activities, and social connectivity features. Their interactive nature stimulates older adults through engaging exercise and interaction, fostering a sense of community and connection [32]. Robots such as Mini and Tangy enhance cognitive abilities in older adults using educational games and imitation learning, aiming to improve memory, cognitive function, and mental agility through innovative stimulation and support [37].

Internet of Things

The Internet of Things (IoT) has revolutionized health care monitoring by using novel sensors and wireless communication methods such as Bluetooth, Wi-Fi, near-field communication, and Zigbee to access and securely store real-time data on vital signs, including temperature, blood pressure, heart rate, and

glucose levels, in the cloud. This advancement facilitates effective device management and real-time data monitoring and offers a secure storage solution, enhancing patient care and treatment outcomes [47]. Furthermore, the integration of multiple sensors into wearable devices, environment monitoring systems, and in-home setups provides pervasive and unobtrusive intelligence to support individuals proactively in their daily lives. These sensors enable continuous monitoring of various parameters, including temperature, humidity, and blood pressure, offering valuable insights for personalized health care interventions [39,42].

IoT-based tools such as smart pillboxes assist individuals in medication management by sending reminders according to prescribed medication plans, ensuring adherence to treatment regimens, and improving medication compliance [39]. Moreover, the implementation of semantic frameworks and advanced data processing techniques enables the collection and analysis of multi-sensor data to determine disease progression and provide timely alerts through real-time monitoring. This approach aids in the early detection of health deterioration, facilitating timely interventions and improved patient outcomes [40]. In addition, the DCARE model used a computational approach to process data from monitoring devices, incorporating content prediction, behavior recognition, and content recognition. This model serves as a comprehensive solution for general monitoring, data processing, and sending alerts to caregivers in case of potential risks or emergencies, enhancing patient safety and care coordination [41].

Virtual, Augmented, and Mixed Reality

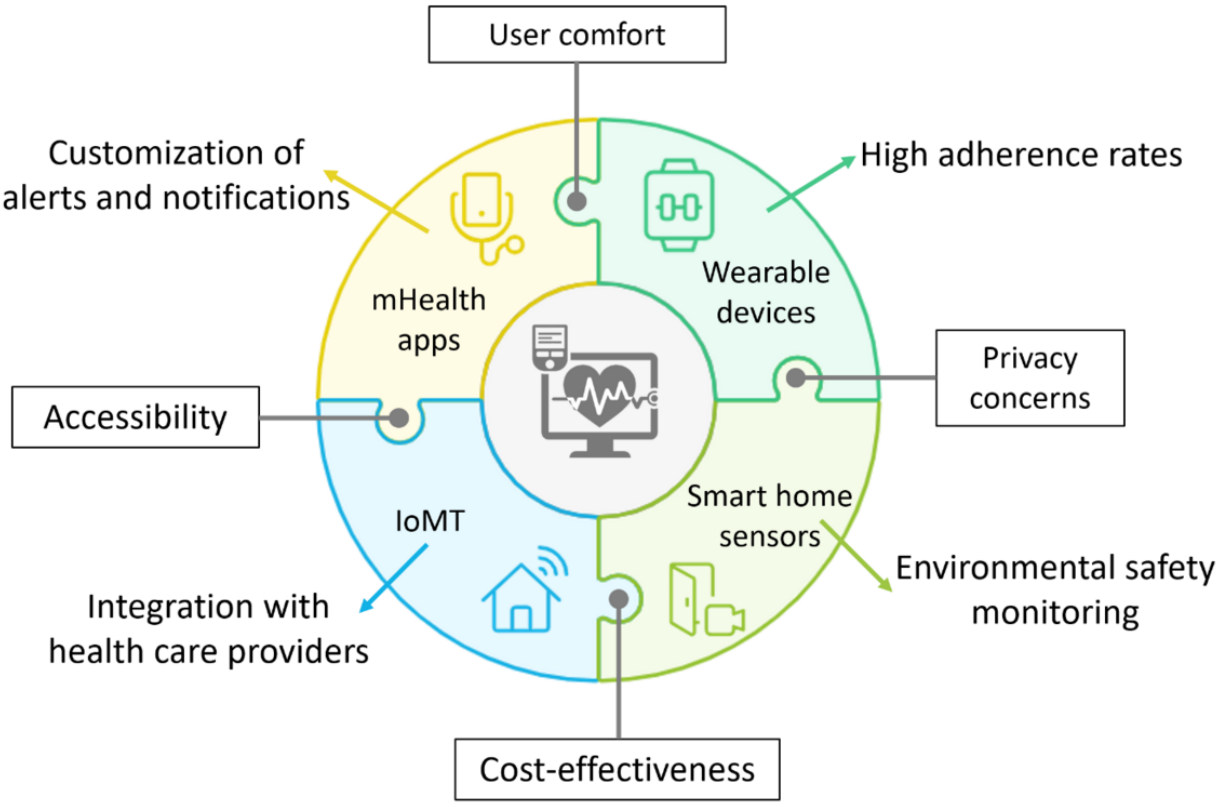
Virtual reality, augmented reality, and mixed reality technologies, facilitated by devices such as headsets, smart glasses, and next-generation smartphones, offer an immersive experience that extends beyond entertainment to include web-based home shopping, leisure activities, and communication, enriching both social interaction and physiological well-being [42]. Smart spectacles equipped with machine learning algorithms for face and object recognition enable personalized video and image playback, assisting patients with dementia and memory loss by stimulating memory recall and cognitive engagement [43]. In addition, GPS and device tracking functionalities, combined with medication reminders via Bluetooth connectivity, empower caregivers to monitor patient locations and medication adherence, ensuring timely interventions and enhancing patient safety [44].

Effectiveness of Remote Monitoring Technologies (RQ 2)

Overview

In addressing RQ 2, which investigated the effectiveness of remote monitoring technologies for individuals with ADRD, a thorough review of previous research papers was conducted. The findings are synthesized and summarized in Multimedia Appendix 4 [14,32,33,36-38,42,45,46,48-52] and Figure 3, providing a detailed comparison of various monitoring methods, their advantages, and their limitations. This comprehensive assessment helps achieve an optimal balance in addressing the needs of both patients and caregivers. The following subsections delineate different aspects pertaining to RQ 2.

Figure 3. Key features and considerations of different remote monitoring technologies. IoMT: Internet of Medical Things; mHealth: mobile health.



Wearable Devices

Wearable devices have emerged as valuable tools for remote monitoring of various health conditions, including cerebrovascular diseases, neurodegenerative diseases, and MCI. Godkin et al [48] investigated the adherence value of wearable devices focusing on 3 key aspects—adherence based on sensor location, time of day, and wear location—and participant and study partner acceptance of continuous multi-sensor wear for a week-long period. The outcomes of this study demonstrated the feasibility of a continuous multi-sensor approach to remote monitoring for individuals living with cerebrovascular diseases and neurodegenerative diseases, showing high adherence rates during both day- and nighttime wear throughout the study period. The aforementioned study also addressed concerns related to device discomfort, including itchiness, catching on clothing during dressing or on bed sheets during sleep, device size, and appearance.

Similarly, a comprehensive investigation conducted by Stavropoulos et al [14] focused on public involvement activities to capture the preferences, priorities, and concerns of people with AD and their caregivers regarding the use of monitoring wearables. Face-to-face discussions with the patient advisory board revealed varying levels of acceptance toward new technologies among older individuals, highlighting concerns regarding the cost of devices, adaption to modern technologies, and potential difficulties in operating technology. Furthermore, Stavropoulos et al [14] used a questionnaire to understand the specific requirements, preferences, barriers, and daily obstacles faced by individuals with MCI, their caregivers, and HCPs in using wearables. The Human Factors and Technology Requirements Questionnaire collected survey responses to various types of questions, such as open-ended, multiple-choice, and closed-ended questions, and the importance of user-centered design for assistive technologies using wearables and highlighted the implications for future research. The aforementioned study provided valuable insights into the perspective of individuals with MCI, caregivers, and HCPs regarding the use of assistive technologies with wearable trackers, emphasizing the need to consider privacy and ethical issues in the technology design and the disadvantages of traditional record-keeping methods.

Smart Homes (Multiple Sensors)

Smart home technologies equipped with multiple sensors have shown promising benefits in remote monitoring, particularly through video monitoring methods. However, there is a growing interest in exploring novel remote monitoring technologies beyond traditional video monitoring, emphasizing the need to diversify research in this area [36]. Despite the potential advantages, there is limited evidence to conclusively determine the acceptance of these remote sensing technologies, indicating a gap in understanding alternative methods [36]. In a comprehensive systematic review, it was highlighted that smart home remote monitoring has the potential to support older persons and caregivers by addressing individual family needs [49]. Nevertheless, critical barriers need to be addressed for successful adoption and implementation, including concerns about the time, training, and technical challenges associated with smart home technologies [49]. These challenges include issues with battery life, unstable wireless connectivity, and incompatibility, which have raised concerns about potential feelings of loneliness and isolation among older people with cognitive impairments [49]. Furthermore, there is a need for more research on alternative remote sensing technologies, such as vibration and acoustics technologies, to better understand their potential applications and effectiveness [36].

Addressing stigma and negative associations with disability through appropriate design and customization of sensing technologies, such as the use of smart fabric sensors, could help mitigate privacy concerns and enhance acceptability [36]. Future research should focus on understanding end-user conditions for acceptability and addressing various barriers, including usability, technical issues, cost concerns, and social implications, to facilitate the broader adoption of smart home technologies in health care settings [36,49].

Multimode Sensor (Machine Learning Approach)

In addressing the challenges of predicting and managing the behavioral and psychological symptoms of dementia (BPSD), multimode sensor technologies coupled with machine learning approaches have emerged as promising tools [45]. These approaches leverage diverse data sources to offer valuable insights into the relative contributions of various factors, such as caregiver-perceived triggers, cognitive and functional status, personality traits, and actigraphy-derived parameters, in predicting BPSD. However, the efficacy of these technologies can be compromised by challenges related to the improper wear of devices, leading to limited data availability for accurate prediction [45]. Despite these obstacles, machine learning models have demonstrated their capability to effectively predict different subsyndromes of BPSD. Such predictive capabilities hold significant potential to inform personalized interventions tailored to individual symptom profiles, thus facilitating timely and targeted care.

Nevertheless, it is essential to acknowledge the existing limitations and gaps in current research [45]. Cho et al [45] emphasize the need for future studies to address these challenges. This includes validating findings in larger and more diverse populations, integrating objective monitoring technologies, and developing more precise prediction algorithms to enhance the accuracy and reliability of these predictive models. Such advancements could significantly contribute to improving the quality of care and support provided to individuals living with dementia and their caregivers.

Patient and Public Involvement Survey

Patient and public involvement surveys emphasize the importance of prioritizing user-friendly design and personalized approaches when developing mHealth apps tailored for individuals with cognitive impairment [50]. Survey findings in the study by Lazarou et al [50] highlighted participants' keen interest in brain games aimed at memory enhancement, showcasing a strong demand for cognitive support tools. In addition, the participants expressed a desire for medication reminders and GPS tracking functionalities within mHealth apps, demonstrating the potential utility of such features in enhancing daily living and emergency response for individuals

with cognitive impairment [50]. Despite these valuable insights, it is important to acknowledge the survey's limitations. The research was conducted among a limited population, potentially introducing self-bias and not capturing the full spectrum of needs and preferences across diverse groups [50]. To address these challenges, future research should aim to explore more diverse populations and incorporate a broader range of user perspectives. This will help in developing more inclusive and effective mHealth interventions that cater to the diverse needs of individuals with cognitive impairment, caregivers, and HCPs.

Internet of Medical Things and Robots

The Internet of Medical Things (IoMT) is revolutionizing the health care landscape by integrating advanced technologies into smart health care systems. In the study by Mathkor et al [52], the authors provided a comprehensive overview of IoMT, highlighting its role in current and future biomedical systems. The study emphasized the integration of technologies such as deep learning, machine learning, blockchain, AI, radio-frequency identification, and Industry 5.0. These technologies hold immense potential in enhancing patient care, improving diagnostic accuracy, and streamlining health care operations. However, the implementation of IoMT systems presents several challenges that need to be addressed for successful adoption. High costs associated with device development, infrastructure setup, and connectivity solutions can pose financial barriers. In addition, many IoMT devices operate on battery power, leading to concerns about battery life and continuous operation. Data quality issues, including noise, biases, and inaccuracies, can also arise due to device limitations or environmental factors. To fully realize the benefits of IoMT, it is crucial to address these challenges and develop robust, scalable, and secure solutions. Companion robots have emerged as potential tools to address the social and emotional needs of older adults, particularly those with dementia.

In the study by Pou-Prom et al [51], the authors conducted an experiment involving 19 participants diagnosed with AD to explore the feasibility of using conversational companion robots. The study revealed that, while some participants enjoyed interacting with the robots, others expressed doubtfulness and reluctance. Technical limitations such as speech recognition challenges, lack of dialogue context, and keyword matching were identified as barriers to seamless communication. Furthermore, the absence of features such as miscommunication and acoustic features hindered the assessment of dementia-related symptoms through robot interactions. The mixed reactions from participants highlighted the importance of ongoing research to refine companion robot technologies. Future work should focus on addressing the identified limitations, exploring the trade-off between autonomous and semiautonomous approaches, and optimizing the design and functionality of these robots to better support the well-being of older adults. Collaborative efforts involving HCPs, engineers, and end users are essential to develop more effective and user-friendly companion robot solutions that can have a meaningful impact on the lives of older adults with dementia.

Cognitive Monitoring for Dementia Care

Advancements in remote monitoring have enabled the integration of cognitive assessment tools to track disease progression and evaluate treatment effectiveness in dementia care. Digital cognitive assessments such as smartphone-based memory and attention tests offer frequent, scalable evaluations that provide real-time insights into cognitive changes. In addition, AI-driven speech and typing analysis can detect subtle linguistic and motor impairments associated with cognitive decline, enabling early intervention and treatment adjustments. Gaze tracking and eye movement analysis further enhance cognitive monitoring by assessing reading patterns, recognition abilities, and attention shifts, which may indicate neurological changes over time. Meanwhile, wearable electroencephalography devices and neurofeedback tools provide continuous, objective measurements of brain activity, allowing researchers and clinicians to assess how therapeutic interventions impact cognitive function. These emerging technologies hold great promise for improving personalized treatment strategies, monitoring drug efficacy, and facilitating early diagnosis, ultimately enhancing the quality of care for individuals with dementia.

Privacy and Security in Remote Monitoring (RQ 3)

Overview

When examining the critical security and privacy considerations involved in remote monitoring technologies, particularly for patients with AD, it is essential to delve into several layers of complexity that encompass data security, privacy regulations, ethical issues, and the unique vulnerabilities of the patient population. Remote monitoring technologies, including smartphones, wearables, and environmental sensors, offer significant benefits for managing AD by enabling continuous health monitoring and early detection of anomalies and providing valuable data for research. However, these benefits come with substantial security and privacy challenges. Different aspects of privacy and security concerns are summarized in Multimedia Appendix 5.

Data Types and Associated Risks

First, the nature of the data collected by these devices—ranging from personal details to medical records, geolocation, and behavioral patterns—makes them highly sensitive. Ensuring data security is paramount as it requires robust encryption methods to protect data both in transit and at rest. Encryption safeguards the data from unauthorized access during transmission and storage. Equally important is implementing strong access control measures, ensuring that only authorized individuals such as health care providers and approved caregivers can access the patient data. Regular security audits and vulnerability assessments are crucial to identifying and addressing potential security gaps, thereby maintaining the integrity of the monitoring systems.

Privacy concerns are equally critical, particularly given the stringent requirements set forth by regulations such as the General Data Protection Regulation (GDPR) in Europe and HIPPA (Health Insurance Portability and Accountability Act) in the United States [58-60]. These regulations mandate that

patients' personal health information be protected and that individuals have control over their data. Adherence to these regulations involves ensuring that patients (or their legal guardians) provide informed consent for data collection and use. This consent process must be clear and comprehensive, detailing what data will be collected, how they will be used, who will have access, and how long the data will be retained. Data minimization principles should be applied, collecting only the necessary data required for the intended purpose, thus reducing the risk of unnecessary exposure. Patients with AD may not fully understand the implications of data collection and their rights regarding data privacy. Therefore, it is imperative to involve caregivers or legal guardians in the consent process to ensure that the patients' rights and preferences are respected. Furthermore, patients with AD are particularly vulnerable to exploitation and abuse. Continuous monitoring can inadvertently infringe on their privacy and autonomy, raising ethical concerns. Balancing the need for constant monitoring to ensure safety with the patients' rights to privacy and dignity is a delicate task. Ethical considerations must guide the implementation of these technologies, ensuring that monitoring practices do not become overly intrusive.

Device and IoT Security

From a technical perspective, the integration of various monitoring devices such as smartphones, wearables, and environmental sensors introduces additional security challenges. Each device and sensor can be a potential entry point for cyberattacks. IoT devices are often targeted due to their typically lower security standards. To counteract these risks, it is crucial to ensure that all devices are secured against tampering and hacking. This involves regular firmware updates to patch vulnerabilities, the use of strong authentication mechanisms to prevent unauthorized device access, and the use of data anonymization techniques to protect patient identity in the event of a data breach [41,61].

Furthermore, the use of advanced technologies such as AI and machine learning in processing and analyzing the collected data presents its own set of risks. While these technologies can enhance the accuracy and efficiency of health monitoring, they also require access to large datasets, which can exacerbate privacy concerns if not handled properly. Ensuring that AI and machine learning models are trained on anonymized data and that they adhere to strict privacy standards is essential to mitigate these risks. Potential attackers pose significant threats to the security of remote monitoring systems. These attackers might attempt to intercept data transmission, gain unauthorized access to monitoring devices, or exploit vulnerabilities to extract sensitive information. Addressing these threats requires a multilayered security approach, including end-to-end encryption, secure communication protocols, and robust incident mechanisms to quickly address and mitigate breaches when they occur [47].

Balancing Safety and Patient Autonomy

While remote monitoring offers substantial benefits for managing AD, it also presents critical security and privacy challenges that must be carefully addressed. Ensuring robust data security through encryption, access controls, and regular audits; complying with privacy regulations; obtaining informed consent; and addressing the unique vulnerabilities of patients with AD are essential components of a secure and ethical monitoring system [62,63].

Balancing safety and patient autonomy in remote monitoring for AD care presents a complex challenge that requires a nuanced approach. While the technology offers substantial benefits, such as continuous health monitoring, early anomaly detection, and timely interventions, it must be implemented thoughtfully to prevent undue intrusion. The need for constant observation can infringe on the patients' sense of independence, raising ethical questions about dignity and consent. To address these concerns, remote monitoring systems should integrate customizable privacy settings that allow caregivers to modify the monitoring level based on the patients' current health status and preferences. Noncritical data anonymization and local data processing using edge computing can significantly reduce the sense of surveillance, ensuring that only essential information is transmitted to central servers [64].

Obtaining informed consent from patients with AD is another critical aspect due to cognitive impairments that often affect their decision-making capabilities. In such cases, involving caregivers or legal guardians becomes essential to uphold the patients' rights while ensuring that their best interests are protected [65,66]. Technologically secure consent management systems can be used to store and track consent records digitally. Leveraging blockchain technology for tamper-proof consent logs can enhance transparency and accountability, making it clear who has agreed to data collection and how the data will be used. This approach ensures compliance with regulations such as the GDPR [60] and HIPAA [58,67], reinforcing trust among patients, caregivers, and health care providers.

Adaptive monitoring systems that support variable levels of oversight can help maintain patient autonomy by shifting from comprehensive tracking to minimal oversight as needed. This adaptability allows caregivers to balance the safety needs with the patients' right to privacy. Moreover, implementing user-controlled privacy features empowers patients and their caregivers to configure data sharing preferences, such as choosing which types of data are collected and how frequently they are transmitted. This flexibility can significantly improve user satisfaction and trust in the monitoring system.

AI integration into monitoring systems should prioritize ethical standards, ensuring that data handling is minimal and transparent. For instance, using federated learning allows algorithms to be trained directly on local devices, avoiding the need to transfer raw data to external servers [68]. This method safeguards privacy while maintaining the system's ability to deliver valuable insights. In addition, explainable AI frameworks can provide justifications for alerts and decisions, ensuring that caregivers understand the reasoning behind the system's actions [69]. Such transparency bolsters confidence in the technology and facilitates informed decision-making.

To enhance safety protocols without compromising privacy, remote monitoring systems should use multilayered privacy protocols that combine encryption, robust user authentication, and differential privacy [70]. By injecting noise into datasets,

differential privacy makes it difficult to trace data back to individual patients while allowing for meaningful analysis. Regular security audits and ethical reviews should also be conducted to ensure that monitoring practices do not become overly invasive. Independent ethics committees can provide oversight, ensuring that the technology aligns with patient-centric values and preserves autonomy.

Maintaining patient dignity is paramount, especially when using advanced monitoring technology. Nonvisual sensors such as microradar [71] systems can detect falls and wandering without capturing images or videos, offering a less invasive alternative to camera-based surveillance. Moreover, patients and caregivers should have the option to customize notification settings, choosing to receive only high-priority alerts to minimize stress. This selective alert system can enhance quality of life by allowing patients to maintain a degree of normalcy and autonomy in their daily lives.

Potential Remote Monitoring Framework for Patients With AD (RQ 4)

Overview

Critical design considerations for a potential remote monitoring framework for patients with AD include ensuring privacy and data security to comply with HIPAA regulations, using nonintrusive monitoring technologies such as microradar sensors to maintain patient dignity, and incorporating AI-driven analytics for accurate detection of falls and wandering. The system should offer real-time alerts to caregivers, be user-friendly for both patients and caregivers, and integrate with existing health care platforms for seamless data sharing. In addition, providing customizable and scalable features can accommodate diverse patient needs and facilitate broader adoption.

The development of our proposed remote monitoring framework was directly informed by the insights gained from our literature review. Our findings revealed that existing monitoring solutions such as wearable sensors and camera-based systems present significant limitations in terms of user compliance, privacy concerns, and real-time responsiveness. Wearable devices require active patient participation, which may not always be feasible for individuals with ADRD, whereas camera-based systems raise ethical and privacy concerns. In response, our framework integrates microradar technology, which allows for the nonintrusive monitoring of falls and wandering behaviors without compromising user privacy. In addition, AI-powered analytics enhance the accuracy of detection by processing microradar data to identify anomalies in movement patterns. To support real-time caregiver intervention, our system uses Web Real-Time Communication (WebRTC) and peer-to-peer (P2P) communication for seamless alerts. By addressing these limitations and leveraging interdisciplinary advancements, our framework offers a comprehensive, privacy-conscious, and scalable solution for the monitoring of patients with ADRD. Different aspects pertaining to RQ 4 are discussed in the following subsections.

Ensuring Patient Safety

Patient safety is the cornerstone of any remote monitoring system for patients with AD. Wearable devices such as smartwatches and fitness trackers can monitor crucial health indicators such as gait, sleep patterns, physical activity, and falls. Environmental sensors placed in the patient's home can track temperature, motion, and other conditions to detect unusual behavior, such as wandering or falls, ensuring that immediate action can be taken [72].

Given the sensitivity of health data, robust security measures are essential. Data from wearable devices and sensors must be encrypted to prevent unauthorized access. Compliance with regulations such as the GDPR and HIPAA is necessary to protect patient privacy. Strong authentication methods and regular software updates can help maintain data integrity and security.

Ease of Use for Caregivers

The system must be user-friendly, as caregivers for individuals with ADRD may not be technologically savvy. A simple, intuitive interface is crucial, allowing for straightforward navigation and clear, actionable alerts to reduce ambiguity and stress. Customizable notifications and automated alerts ensure that caregivers are informed of critical issues without needing to constantly monitor the system while minimizing false alarms to avoid unnecessary stress. In addition, providing educational resources within the system helps caregivers understand how to best use the monitoring tools and improve their caregiving skills. This approach empowers caregivers, enhances efficiency, and improves patient safety by facilitating timely interventions [73,74].

Ethical Considerations

Ethical considerations for a potential remote monitoring framework for patients with AD include ensuring informed consent and respecting patient autonomy. Considering cognitive impairments, obtaining informed consent must involve patients as much as possible as well as legal representatives or caregivers when necessary to ensure understanding and voluntary participation. Privacy and confidentiality are paramount, requiring compliance with data protection regulations and secure data handling practices. The system should be minimally intrusive, using discreet technologies to balance safety with patient dignity and avoiding constant visible monitoring that could lead to a loss of independence. Transparency in communication about the system's operation and data use, as well as providing feedback mechanisms, builds trust and addresses concerns. Ethical design principles involving user-centered approaches and bias mitigation in AI, alongside regular ethical reviews and oversight by ethics committees, ensure that the system remains respectful and patient centered while enhancing safety [75,76].

Compatibility With Existing Systems

To be effective, the remote monitoring system for patients with AD must integrate seamlessly with existing health care infrastructure, such as electronic health records (EHRs) [77]. This integration ensures that all relevant health data from the monitoring system are accurately and efficiently captured within the broader health care framework. Using standardized protocols

and application programming interfaces is critical for facilitating data sharing and interoperability [78]. These standards, such as Health Level Seven Fast Healthcare Interoperability Resources, ensure that data from the monitoring system can be easily and securely exchanged with different EHR systems and other health IT platforms [79]. Seamless integration allows physicians, specialists, and other clinicians to access comprehensive, up-to-date information about patients, promoting coordinated care. For instance, if a remote monitoring system detects a fall or wandering incident, this information can be automatically logged in the patient's EHR, alerting HCPs in real time. This ensures that primary care physicians, specialists, and hospital staff have a unified view of the patient's health status and recent events, enabling more informed decision-making and timely interventions.

Moreover, compatibility with existing systems helps in reducing duplication of efforts and minimizes the risk of errors. HCPs can avoid manually entering data, which not only saves time but also reduces the likelihood of transcription errors. By ensuring that the remote monitoring data are automatically integrated into EHRs, the system enhances the accuracy and reliability of patient records. In addition to technical interoperability, organizational processes and workflows must be considered to ensure smooth integration. Training for physicians, nurses, and other HCPs on how to use the new system and interpret its data within the context of existing EHRs is essential. Developing clear protocols for how and when data from the monitoring system should be used in patient care can help ensure that the system is used effectively. By enabling seamless data sharing and integration, the system supports a more coordinated and comprehensive approach to health care, ultimately improving outcomes for patients with AD and reducing the burden on caregivers and medical professionals.

Scalability and Flexibility

A scalable system is essential for accommodating the increasing number of patients with AD who require care. As demand for remote monitoring solutions grows, the system must expand without compromising performance or reliability. This need involves designing the system architecture to support a larger user base and additional data processing and storage capabilities. Scalable infrastructure ensures that the system can handle a surge in the number of patients and caregivers, allowing health care providers to maintain high standards of care even as the user population increases. Moreover, incorporating standardized interfaces and protocols such as application programming interfaces and interoperability standards ensures that the system can easily connect with new devices and technologies, facilitating continuous improvement and innovation [80].

Flexibility in the remote monitoring system is equally important. The system should have a modular architecture that allows for the easy integration of new sensors and technologies as they become available. This modular approach ensures that the system can adapt to technological advancements and incorporate innovative solutions without requiring a complete redesign. It should also be configurable to meet the specific requirements of various health care settings, such as in-home care, assisted living facilities, or hospitals, providing tailored solutions that address the unique challenges of each environment. The system's scalability and flexibility ensure that it adapts to evolving health care needs, enhancing patient care and supporting the growth of remote monitoring technologies [80].

Technology Components

Wearable technology is essential for health monitoring, allowing for the real-time tracking of multiple health variables such as heart rate, physical activity, sleep habits, and location [56,72]. These devices help monitor the patient's daily routines and identify emergencies such as falls. Environmental sensors also enhance the patient's living environment by monitoring door activity, motion, and room temperature to detect unusual activity. For instance, if a patient leaves the house at an unusual time, an alert can be issued to the caregiver. Smartphone apps enable caregivers to easily monitor and receive notifications, enter observations, create personalized alerts, and access learning materials. Continuous connectivity ensures that caregivers are always informed about the patient's status.

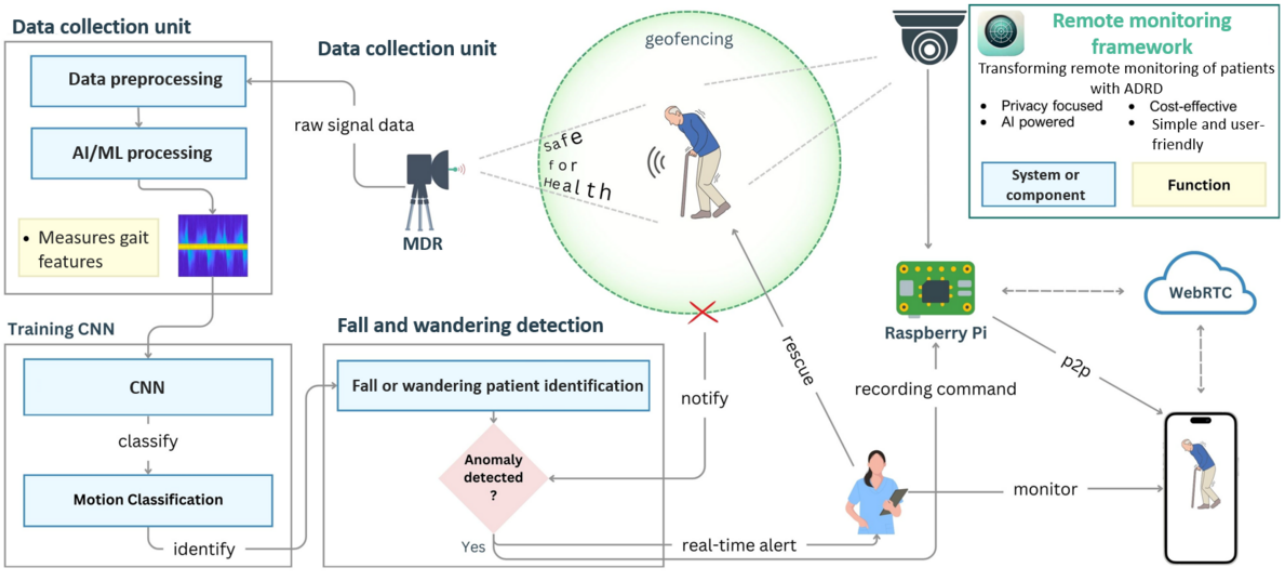
Implementing AI and machine learning algorithms to analyze data collected from wearables and other devices allows for the identification of trends indicating potential health issues. These predictive analytics enable caregivers to anticipate problems and take preventive measures before they become critical, thus improving patient care and safety. Moreover, an IoT infrastructure is essential for ensuring seamless data transfer between all connected devices, with cloud computing offering scalable processing and storage capacity and edge computing minimizing delays by processing data locally. Effective data management is crucial for making sense of the vast amounts of data generated by the system, allowing for trend analysis and informed clinical decisions. Advanced analytics techniques can also provide personalized patient care. In addition, robust security technologies such as intrusion detection systems, firewalls, and encryption methods are vital to protect the system from threats, with frequent maintenance and updates ensuring ongoing security

Comprehensive Framework

Overview

On the basis of the discussions in the previous subsections and the findings of previous studies, we propose a comprehensive remote monitoring framework, as shown in Figure 4, by incorporating the use of wearable technologies, sensors, AI-powered analytics, and a safe IoT infrastructure. The elements in the schematic framework are intricately connected and together create a comprehensive system for the remote monitoring of patients with AD. The framework is designed with a strong focus on health and safety. The use of microradar technology ensures that monitoring is nonintrusive and safe for patients, reducing the stress associated with traditional monitoring systems. This framework not only enhances patient safety but also alleviates the burden on caregivers by providing reliable, real-time support and intervention tools. The following is a detailed discussion of how each point contributes to the overall monitoring framework.

Figure 4. Schematic framework for remote monitoring of patients with Alzheimer disease. ADRD: Alzheimer disease and related dementias; AI: artificial intelligence; CNN: convolutional neural network; MDR: micro-Doppler radar; ML: machine learning; P2P: peer to peer; WebRTC: Web Real-Time Communication.



Data Collection

The data collection unit is the central component of the monitoring system and is responsible for gathering and processing the data required to track the patient's movements and health status effectively. At its core, the unit uses micro-Doppler radar technology to capture raw signal data by continuously monitoring the patient's movements in a nonintrusive manner [81]. This technology is crucial as it does not require the patient to wear any devices or interact with the system, thereby preserving the patient's privacy and minimizing any stress or discomfort that might arise from traditional monitoring systems. The raw data collected by the micro-Doppler radar undergoes a critical preprocessing stage in which essential gait features are extracted. This step is vital in filtering out noise and isolating the most relevant aspects of the patient's movements that could signal potential risks. Once preprocessed, the data are fed into advanced AI and machine learning algorithms, which analyze the extracted features to identify patterns in the patient's movements. This sophisticated processing enables the system to distinguish between normal daily activities and potentially hazardous situations such as falls or wandering, thereby enhancing the system's ability to ensure the patient's safety while maintaining their dignity and comfort.

Data Processing

The data processing component plays a critical role in enhancing the system's ability to accurately classify various patient movements, which is fundamental to ensuring effective monitoring and maintaining the safety of individuals with ADRD. The process begins with preprocessed data being fed into a convolutional neural network [82], a powerful AI model specifically designed to recognize and analyze patterns within the input data. This network is trained to classify different types of motions, enabling it to distinguish between normal daily activities such as walking or sitting and potentially hazardous incidents that may require immediate intervention, such as falls or wandering outside a designated safe zone. The convolutional neural network's ability to make these distinctions is vital to the system's overall accuracy and reliability as it allows for the early detection of unusual or dangerous behaviors. This, in turn, facilitates timely alerts and responses, ensuring that caregivers can intervene quickly and appropriately to protect the patient's well-being. Through this advanced data processing approach, the system not only enhances patient safety but also provides peace of mind to caregivers through the knowledge that the system is capable of accurately monitoring and responding to the specific needs of patients with ADRD.

Fall and Wandering Detection

The fall and wandering detection component is a critical part of the monitoring system, designed to ensure the continuous safety of patients by closely tracking their movements and identifying any signs of falling or wandering outside a predefined safe area. The system uses a geofencing approach in which a *defined* boundary or safe zone is established around the patient's environment. This safe zone is carefully configured based on the patient's needs and typical movement patterns. The system continuously monitors the patient's movements within this boundary using sophisticated pattern identification techniques. By analyzing these movement patterns in real time, the system can quickly detect any deviations from the norm. For example, if the patient crosses this virtual boundary or exhibits movements that suggest a fall, such as a sudden loss of balance or an abrupt change in posture, the system immediately flags this behavior as an anomaly. This continuous, vigilant monitoring is essential for identifying potential risks before they can lead to serious harm.

Once an anomaly is detected, the system's response should be both swift and automated. The anomaly detection process should be finely tuned to recognize specific triggers, such as the patient moving outside the safe zone or experiencing a fall. When such an event is identified, the system automatically triggers an alert, ensuring that caregivers are notified in real time. These alerts should be designed to reach caregivers through various channels,

such as mobile notifications or direct messages, providing them with immediate information about the patient's situation. This real-time notification system is crucial for enabling caregivers to respond promptly and appropriately to any incidents, whether it involves assisting the patient after a fall or guiding them back to safety after wandering. The combination of continuous monitoring and instantaneous alerting forms a robust safety net, significantly reducing the risks associated with falls and wandering in patients with ADRD.

Caregiver Notification and Support

The caregiver notification and support component is designed to empower caregivers with the tools and information they need to effectively monitor and respond to the needs of patients in real time. Central to this component are real-time alerts, which are instantly sent to caregivers' devices, ensuring that they are always aware of the patient's status. These immediate notifications are particularly crucial in emergency situations, such as when a fall or wandering event occurs. By receiving alerts the moment an anomaly is detected, caregivers can intervene promptly, providing necessary assistance or preventing further harm to the patient. This real-time communication is key to maintaining the safety and well-being of patients with ADRD.

In addition, the system is equipped with a recording command and rescue feature facilitated by a Raspberry Pi module. This module plays a vital role by capturing and recording the event as soon as an anomaly is detected. The recorded data are automatically sent to caregivers, offering them detailed insights into the incident. This documentation not only aids in immediate rescue actions but also serves as a valuable resource for later analysis. By reviewing recorded incidents, caregivers and medical professionals can better understand the circumstances leading up to the event, which is essential for medical evaluations and refining future care strategies. The ability to document and analyze these events ensures that care can be continuously improved, ultimately enhancing the quality of life of both patients and caregivers.

WebRTC and P2P Communication

The framework integrates advanced communication technologies to ensure seamless monitoring and interaction between caregivers and the system, providing a reliable and immediate connection regardless of location. A key component of this system is the use of WebRTC technology, which facilitates real-time P2P communication [83,84]. This technology allows caregivers to monitor the patient from their smartphones or other devices with ease and without delay, ensuring that they remain connected to the patient's status at all times. The use of WebRTC ensures that the communication is not only immediate but also direct, reducing latency and enhancing the responsiveness of the system.

This capability significantly enhances patient safety as it allows caregivers to maintain a continuous, real-time connection with the system, enabling them to respond quickly to any situation that may arise. Whether it is monitoring routine activities or responding to alerts triggered by the system, WebRTC ensures that caregivers can provide timely and effective interventions. The seamless monitoring facilitated by this technology is crucial for reinforcing the safety net around patients with ADRD, giving caregivers the confidence and tools they need to manage their loved ones' care effectively, no matter where they are.

Conclusions

This study highlights the revolutionary potential of remote monitoring technologies in AD care, emphasizing the critical balance among innovation, practicality, and empathy. The comprehensive analysis revealed that, while wearable devices and ambient sensors significantly enhance the ability to track and support patients with AD, they also introduce challenges related to privacy, data security, and user acceptance. The findings stress the necessity of involving patients and caregivers in the design process to ensure that these technologies meet their specific needs and regulatory requirements, which is crucial to the successful implementation of these technologies.

The proposed remote monitoring framework integrates cutting-edge microradar technology with AI-powered processing to create a nonintrusive, privacy-conscious solution for monitoring patients with AD and dementia. By focusing on real-time data processing, caregiver notification, and seamless communication, the system not only enhances patient safety but also provides critical support to caregivers, ultimately improving the quality of care for individuals with ADRD. Future research can focus on refining these systems to enhance their usability, accessibility, and security, thereby improving the overall caregiving experience and patient well-being. By advancing remote monitoring practices, we can create a supportive environment that alleviates caregiver burdens and promotes a higher quality of life for individuals living with AD.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 31 KB - aging_v8i1e69175_app1.docx]

Multimedia Appendix 2
Remote monitoring technologies using wearable and portable device–based sensors.
[DOCX File , 45 KB - aging_v8i1e69175_app2.docx]

Multimedia Appendix 3
Remote monitoring technologies using nonwearable home sensors.
[DOCX File , 44 KB - aging_v8i1e69175_app3.docx]

Multimedia Appendix 4
Remote monitoring technologies for dementia care and their role in cognitive assessment.
[DOCX File , 49 KB - aging_v8i1e69175_app4.docx]

Multimedia Appendix 5
Different privacy and security considerations in remote monitoring.
[DOCX File , 14 KB - aging_v8i1e69175_app5.docx]

References

1. Cipriani G, Danti S, Picchi L, Nuti A, Fiorino MD. Daily functioning and dementia. Dement Neuropsychol 2020;14(2):93-102 [FREE Full text] [doi: 10.1590/1980-57642020dn14-020001] [Medline: 32595877]
2. Tatulian SA. Challenges and hopes for Alzheimer's disease. Drug Discov Today 2022 Apr;27(4):1027-1043 [FREE Full text] [doi: 10.1016/j.drudis.2022.01.016] [Medline: 35121174]
3. Breijyeh Z, Karaman R. Comprehensive review on Alzheimer's disease: causes and treatment. Molecules 2020 Dec 08;25(24):5789 [FREE Full text] [doi: 10.3390/molecules25245789] [Medline: 33302541]
4. Abeysinghe AA, Deshapriya RD, Udawatte C. Alzheimer's disease; a review of the pathophysiological basis and therapeutic interventions. Life Sci 2020 Sep 01;256:117996 [FREE Full text] [doi: 10.1016/j.lfs.2020.117996] [Medline: 32585249]
5. Lei P, Ayton S, Bush AI. The essential elements of Alzheimer’s disease. J Biol Chem 2021;296:100105 [FREE Full text] [doi: 10.1074/jbc.rev120.008207]
6. Castellani RJ, Rolston RK, Smith MA. Alzheimer disease. Dis Mon 2010 Sep;56(9):484-546 [FREE Full text] [doi: 10.1016/j.disamonth.2010.06.001] [Medline: 20831921]
7. Shin JY, Habermann B. Caregivers of adults living with Alzheimer's disease or dementia in 2020: a secondary analysis. J Gerontol Nurs 2022 Sep;48(9):15-25 [FREE Full text] [doi: 10.3928/00989134-20220805-02] [Medline: 36007215]
8. Reinhard SC, Given B, Petlick NH, Bemis A. Supporting family caregivers in providing care. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville, MD: Agency for Healthcare Research and Quality (US); 2008.
9. Culberson JW, Kopel J, Sehar U, Reddy PH. Urgent needs of caregiving in ageing populations with Alzheimer's disease and other chronic conditions: support our loved ones. Ageing Res Rev 2023 Sep;90:102001 [FREE Full text] [doi: 10.1016/j.arr.2023.102001] [Medline: 37414157]
10. Koca E, Taşkapilioğlu Ö, Bakar M. Caregiver burden in different stages of Alzheimer's disease. Noro Psikiyatr Ars 2017 Mar;54(1):82-86 [FREE Full text] [doi: 10.5152/npa.2017.11304] [Medline: 28566965]
11. Manzini CS, do Vale FA. Emotional disorders evidenced by family caregivers of older people with Alzheimer's disease. Dement Neuropsychol 2020;14(1):56-61 [FREE Full text] [doi: 10.1590/1980-57642020dn14-010009] [Medline: 32206199]
12. Liao X, Huang Y, Zhang Z, Zhong S, Xie G, Wang L, et al. Factors associated with health-related quality of life among family caregivers of people with Alzheimer's disease. Psychogeriatrics 2020 Jul;20(4):398-405 [FREE Full text] [doi: 10.1111/psyg.12528] [Medline: 32027776]
13. Lott SA, Streel E, Bachman S, Bode K, Dyer J, Fitzer-Attas C, et al. Digital health technologies for Alzheimer's disease and related dementias: initial results from a landscape analysis and community collaborative effort. J Prev Alzheimers Dis 2024;11(5):1480-1489. [doi: 10.14283/jpad.2024.103] [Medline: 39350395]
14. Stavropoulos TG, Lazarou I, Diaz A, Gove D, Georges J, Manyakov NV, et al. Wearable devices for assessing function in Alzheimer's disease: a European public involvement activity about the features and preferences of patients and caregivers. Front Aging Neurosci 2021 Apr 12;13:643135 [FREE Full text] [doi: 10.3389/fnagi.2021.643135] [Medline: 33912025]
15. Dieffenderfer J, Brewer A, Noonan MA, Smith M, Eichenlaub E, Haley KL. A wearable system for continuous monitoring and assessment of speech, gait, and cognitive decline for early diagnosis of ADRD. In: Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2023 Presented at: EMBC 2023; July 24-27, 2023; Sydney, Australia. [doi: 10.1109/embc40787.2023.10339986]
16. Chandrasekaran R, Sadiq T M, Moustakas E. Usage trends and data sharing practices of healthcare wearable devices among US adults: cross-sectional study. J Med Internet Res 2025 Feb 21;27:e63879 [FREE Full text] [doi: 10.2196/63879] [Medline: 39982763]
17. Chung J, Pretzer-Aboff I, Parsons P, Falls K, Bulut E. Using a device-free wi-fi sensing system to assess daily activities and mobility in low-income older adults: protocol for a feasibility study. JMIR Res Protoc 2024 Nov 12;13:e53447 [FREE Full text] [doi: 10.2196/53447] [Medline: 39531268]

18. Tiersen F, Batey P, Harrison MJ, Naar L, Serban AI, Daniels SJ, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047 [FREE Full text] [doi: 10.2196/27047] [Medline: 34383672]
19. Hackett K, Xu S, McKniff M, Paglia L, Barnett I, Giovannetti T. Mobility-based smartphone digital phenotypes for unobtrusively capturing everyday cognition, mood, and community life-space in older adults: feasibility, acceptability, and preliminary validity study. JMIR Hum Factors 2024 Nov 22;11:e59974 [FREE Full text] [doi: 10.2196/59974] [Medline: 39576984]
20. Peterson CM, Mikal JP, McCarron HR, Finlay JM, Mitchell LL, Gaugler JE. The feasibility and utility of a personal health record for persons with dementia and their family caregivers for web-based care coordination: mixed methods study. JMIR Aging 2020 Jun 26;3(1):e17769 [FREE Full text] [doi: 10.2196/17769] [Medline: 32589158]
21. Alhammad N, Alajlani M, Abd-Alrazaq A, Epiphaniou G, Arvanitis T. Patients' perspectives on the data confidentiality, privacy, and security of mHealth apps: systematic review. J Med Internet Res 2024 May 31;26:e50715 [FREE Full text] [doi: 10.2196/50715] [Medline: 38820572]
22. Deng C, Shen N, Li G, Zhang K, Yang S. Digital isolation and dementia risk in older adults: longitudinal cohort study. J Med Internet Res 2025 Feb 19;27:e65379 [FREE Full text] [doi: 10.2196/65379] [Medline: 39969956]
23. Davis BH, Tamblin SP, Troutman-Jordan M. Expanding social interaction through online technology: using social media in vascular dementia. In: Huang L, Davis B, editors. Language, Aging and Society. Cham, Switzerland: Palgrave Macmillan; 2024.
24. Zaslavsky O, Kaneshiro J, Chu F, Teng A, Domoto-Reilly K, Chen AT. Virtual intervention for caregivers of persons with Lewy body dementia: pilot quasi-experimental single-arm study. JMIR Form Res 2022 Jul 29;6(7):e37108 [FREE Full text] [doi: 10.2196/37108] [Medline: 35904843]
25. Paolillo EW, Casaletto K, Clark A, Taylor J, Heuer H, Wise A, et al. Examining associations between smartphone use and clinical severity in frontotemporal dementia: proof-of-concept study. JMIR Aging 2024 Jun 26;7:e52831 [FREE Full text] [doi: 10.2196/52831] [Medline: 38922667]
26. Infographic: understanding different types of dementia. National Institutes of Health National Institute on Aging. URL: https://www.nia.nih.gov/health/alzheimers-and-dementia/understanding-different-types-dementia [accessed 2025-02-23]
27. Kalaria R. Similarities between Alzheimer's disease and vascular dementia. J Neurol Sci 2002 Nov 15;203-204:29-34 [FREE Full text] [doi: 10.1016/s0022-510x(02)00256-3] [Medline: 12417353]
28. Deng Z, Guo L, Chen X, Wu W. Smart wearable systems for health monitoring. Sensors (Basel) 2023 Feb 23;23(5):2479 [FREE Full text] [doi: 10.3390/s23052479] [Medline: 36904682]
29. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Hosseinzadeh S, Sajadi Tabar SS. Smartwatches in healthcare medicine: assistance and monitoring; a scoping review. BMC Med Inform Decis Mak 2023 Nov 03;23(1):248 [FREE Full text] [doi: 10.1186/s12911-023-02350-w] [Medline: 37924029]
30. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: an R package and shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. Campbell Syst Rev 2022 Mar 27;18(2):e1230 [FREE Full text] [doi: 10.1002/cl2.1230] [Medline: 36911350]
31. Page MJ, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
32. Popp Z, Low S, Igwe A, Rahman MS, Kim M, Khan R, et al. Shifting from active to passive monitoring of Alzheimer disease: the state of the research. J Am Heart Assoc 2024 Jan 16;13(2):e031247 [FREE Full text] [doi: 10.1161/JAHA.123.031247] [Medline: 38226518]
33. Fares N, Sherratt RS, Elhajj IH. Directing and orienting ICT healthcare solutions to address the needs of the aging population. Healthcare (Basel) 2021 Feb 02;9(2):147 [FREE Full text] [doi: 10.3390/healthcare9020147] [Medline: 33540510]
34. Anghel I, Cioara T, Moldovan D, Antal M, Pop CD, Salomie I, et al. Smart environments and social robots for age-friendly integrated care services. Int J Environ Res Public Health 2020 May 27;17(11):3801 [FREE Full text] [doi: 10.3390/ijerph17113801] [Medline: 32471108]
35. Chouraki A, Tournant J, Arnal P, Pépin JL, Bailly S. Objective multi-night sleep monitoring at home: variability of sleep parameters between nights and implications for the reliability of sleep assessment in clinical trials. Sleep 2023 May 10;46(5):zsac319 [FREE Full text] [doi: 10.1093/sleep/zsac319] [Medline: 36583300]
36. Tong K, Attenborough K, Sharp D, Taherzadeh S, Deepak-Gopinath M, Vseteckova J. Acceptability of remote monitoring in assisted living/smart homes in the United Kingdom and associated use of sounds and vibrations—a systematic review. Appl Sci 2024 Jan 19;14(2):843 [FREE Full text] [doi: 10.3390/app14020843]
37. Read E, Woolsey C, Donelle L, Weeks L, Chinho N. Passive remote monitoring and aging in place: a scoping review. Can J Aging 2022 Aug 01;42(1):20-32 [FREE Full text] [doi: 10.1017/s0714980822000198]
38. Shiwani T, Relton S, Evans R, Kale A, Heaven A, Clegg A, Ageing Data Research Collaborative (Geridata) AI group, et al. New horizons in artificial intelligence in the healthcare of older people. Age Ageing 2023 Dec 01;52(12):afad219 [FREE Full text] [doi: 10.1093/ageing/afad219] [Medline: 38124256]

39. Purohit P, Khanpara P, Patel U, Kathiria P. IoT based ambient assisted living technologies for healthcare: concepts and design challenges. In: Proceedings of the 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud). 2022 Presented at: I-SMAC 2022; November 10-12, 2022; Dharan, Nepal. [doi: 10.1109/i-smac55078.2022.9987375]
40. Giannios G, Mpaltadoros L, Alepopoulos V, Grammatikopoulou M, Stavropoulos TG, Nikolopoulos S, et al. A semantic framework to detect problems in activities of daily living monitored through smart home sensors. Sensors (Basel) 2024 Feb 08;24(4):1107 [FREE Full text] [doi: 10.3390/s24041107] [Medline: 38400265]
41. Machado SD, Tavares JE, Martins MG, Barbosa JL, González GV, Leithardt VR. Ambient intelligence based on IoT for assisting people with Alzheimer's disease through context histories. Electronics 2021 May 25;10(11):1260 [FREE Full text] [doi: 10.3390/electronics10111260]
42. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020 Aug 11;3(2):e17286 [FREE Full text] [doi: 10.2196/17286] [Medline: 32780020]
43. Patil S, Shobha T, Kumari JJ, Khanum R, Anjum R, Manjunath S. Medical assistance for Alzheimer's disease using smart specs. In: Proceedings of the 4th International Conference on Circuits, Control, Communication and Computing. 2022 Presented at: I4C 2022; December 21-23, 2022; Bangalore, India URL: https://doi.org/10.1109/I4C57141.2022.10057684 [doi: 10.1109/i4c57141.2022.10057684]
44. Amato F, Crovari P, Masciadri A, Bianchi S, Pasquarelli MG, Toldo M, et al. CLONE: a promising system for the remote monitoring of Alzheimer's patients: an experimentation with a wearable device in a village for Alzheimer's care. In: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good. 2018 Presented at: Goodtechs '18; November 28-30, 2018; Bologna, Italy. [doi: 10.1145/3284869.3284906]
45. Cho E, Kim S, Heo SJ, Shin J, Hwang S, Kwon E, et al. Machine learning-based predictive models for the occurrence of behavioral and psychological symptoms of dementia: model development and validation. Sci Rep 2023 May 18;13(1):8073 [FREE Full text] [doi: 10.1038/s41598-023-35194-5] [Medline: 37202454]
46. Sapci AH, Sapci HA. Innovative assisted living tools, remote monitoring technologies, artificial intelligence-driven solutions, and robotic systems for aging societies: systematic review. JMIR Aging 2019 Nov 29;2(2):e15429 [FREE Full text] [doi: 10.2196/15429] [Medline: 31782740]
47. Li C, Wang J, Wang S, Zhang Y. A review of IoT applications in healthcare. Neurocomputing 2024 Jan 14;565:127017 [FREE Full text] [doi: 10.1016/j.neucom.2023.127017]
48. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
49. Tian YJ, Felber NA, Pageau F, Schwab DR, Wangmo T. Benefits and barriers associated with the use of smart home health technologies in the care of older persons: a systematic review. BMC Geriatr 2024 Feb 14;24(1):152 [FREE Full text] [doi: 10.1186/s12877-024-04702-1] [Medline: 38355464]
50. Lazarou I, Stavropoulos TG, Mpaltadoros L, Nikolopoulos S, Koumanakos G, Tsolaki M, et al. Human factors and requirements of people with cognitive impairment, their caregivers, and healthcare professionals for mHealth apps including reminders, games, and geolocation tracking: a survey-questionnaire study. J Alzheimers Dis Rep 2021 Jun 11;5(1):497-513 [FREE Full text] [doi: 10.3233/ADR-201001] [Medline: 34368634]
51. Pou-Prom C, Raimondo S, Rudzicz F. A conversational robot for older adults with Alzheimer's disease. ACM Trans Hum Rob Interact 2020 May 31;9(3):1-25 [FREE Full text] [doi: 10.1145/3380785]
52. Mathkor DM, Mathkor N, Bassfar Z, Bantun F, Slama P, Ahmad F, et al. Multirole of the internet of medical things (IoMT) in biomedical systems for managing smart healthcare systems: an overview of current and future innovative trends. J Infect Public Health 2024 Apr;17(4):559-572 [FREE Full text] [doi: 10.1016/j.jiph.2024.01.013] [Medline: 38367570]
53. Sushames A, Edwards A, Thompson F, McDermott R, Gebel K. Validity and reliability of Fitbit flex for step count, moderate to vigorous physical activity and activity energy expenditure. PLoS One 2016 Sep 02;11(9):e0161224 [FREE Full text] [doi: 10.1371/journal.pone.0161224] [Medline: 27589592]
54. Mantua J, Gravel N, Spencer RM. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. Sensors (Basel) 2016 May 05;16(5):646 [FREE Full text] [doi: 10.3390/s16050646] [Medline: 27164110]
55. Petrucci GJ, Freedson P, Masteller B, Cox M, Staudenmayer J, Sirard J. Sensitivity of the Misfit Shine™ to detect changes in laboratory-based and free-living physical activity. J Meas Phys Behav 2018 Mar;1(1):18-25 [FREE Full text] [doi: 10.1123/JMPB.2017-0006]
56. Murakami H, Kawakami R, Nakae S, Yamada Y, Nakata Y, Ohkawara K, et al. Accuracy of 12 wearable devices for estimating physical activity energy expenditure using a metabolic chamber and the doubly labeled water method: validation study. JMIR Mhealth Uhealth 2019 Aug 02;7(8):e13938 [FREE Full text] [doi: 10.2196/13938] [Medline: 31376273]
57. Vlami K, Pantelidi K, Dalamaga M, Karagianni F, Theodoropoulos K, Papiris S, et al. Psoriatic insomnia: a subjective and objective sleep evaluation. Acta Derm Venereol 2023 Feb 14;103:adv00864 [FREE Full text] [doi: 10.2340/actadv.v103.4507] [Medline: 36786739]

58. Oakley A. HIPAA, HIPPA, or HIPPO: what really is the heath insurance portability and accountability act? Biotechnol Law Rep 2023 Dec 13;42(6):306-318 [FREE Full text] [doi: 10.1089/blr.2023.29329.aso]
59. Bradford L, Aboy M, Liddell K. International transfers of health data between the EU and USA: a sector-specific approach for the USA to ensure an 'adequate' level of protection. J Law Biosci 2020 Oct 15;7(1):lsaa055 [FREE Full text] [doi: 10.1093/jlb/lsaa055] [Medline: 34221424]
60. Zaeem RN, Barber KS. The effect of the GDPR on privacy policies: recent progress and future promise. ACM Trans Manage Inf Syst 2020 Dec 08;12(1):1-20 [FREE Full text] [doi: 10.1145/3389685]
61. Gochoo M, Alnajjar F, Tan TH, Khalid S. Towards privacy-preserved aging in place: a systematic review. Sensors (Basel) 2021 Apr 28;21(9):3082 [FREE Full text] [doi: 10.3390/s21093082] [Medline: 33925161]
62. Uddin R, Koo I. Real-time remote patient monitoring: a review of biosensors integrated with multi-hop IoT systems via cloud connectivity. Appl Sci 2024 Feb 25;14(5):1876 [FREE Full text] [doi: 10.3390/app14051876]
63. Shojaei P, Vlahu-Gjorgievska E, Chow YW. Security and privacy of technologies in health information systems: a systematic literature review. Computers 2024 Jan 31;13(2):41 [FREE Full text] [doi: 10.3390/computers13020041]
64. Krishnamoorthy S, Dua A, Gupta S. Role of emerging technologies in future IoT-driven Healthcare 4.0 technologies: a survey, current challenges and future directions. J Ambient Intell Hum Comput 2021 May 12;14:361-407 [FREE Full text] [doi: 10.1007/s12652-021-03302-w]
65. Daly TP. Informing consent to antibodies in Alzheimer's disease. BMJ 2023 Oct 11;383:2350. [doi: 10.1136/bmj.p2350] [Medline: 37821123]
66. Gao SS, Chu CH, Young FY. Oral health and care for elderly people with Alzheimer's disease. Int J Environ Res Public Health 2020 Aug 07;17(16):5713 [FREE Full text] [doi: 10.3390/ijerph17165713] [Medline: 32784777]
67. Nass SJ, Levit LA, Gostin LO. HIPAA, the privacy rule, and its application to health research. In: Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research. Washington, DC: National Academies Press (US); 2009.
68. Li L, Fan Y, Tse M, Lin KY. A review of applications in federated learning. Comput Ind Eng 2020 Nov;149:106854. [doi: 10.1016/j.cie.2020.106854]
69. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, et al. Explainable AI (XAI): core ideas, techniques, and solutions. ACM Comput Surv 2023 Jan 16;55(9):1-33. [doi: 10.1145/3561048]
70. Mahmood Z, Jusas V. Blockchain-enabled: multi-layered security federated learning platform for preserving data privacy. Electronics 2022 May 19;11(10):1624. [doi: 10.3390/electronics11101624]
71. Feng X, Shan Z, Zhao Z, Xu Z, Zhang T, Zhou Z, et al. Millimeter-wave radar monitoring for elder’s fall based on multi-view parameter fusion estimation and recognition. Remote Sens 2023 Apr 16;15(8):2101. [doi: 10.3390/rs15082101]
72. Chakrabarti S, Biswas N, Jones LD, Kesari S, Ashili S. Smart consumer wearables as digital diagnostic tools: a review. Diagnostics (Basel) 2022 Aug 31;12(9):2110 [FREE Full text] [doi: 10.3390/diagnostics12092110] [Medline: 36140511]
73. Hussein T, Chauhan PK, Dalmer NK, Rudzicz F, Boger J. Exploring interface design to support caregivers' needs and feelings of trust in online content. J Rehabil Assist Technol Eng 2020 Dec 08;7:2055668320968482 [FREE Full text] [doi: 10.1177/2055668320968482] [Medline: 33343921]
74. Xie B, Tao C, Li J, Hilsabeck RC, Aguirre A. Artificial intelligence for caregivers of persons with Alzheimer's disease and related dementias: systematic literature review. JMIR Med Inform 2020 Aug 20;8(8):e18189 [FREE Full text] [doi: 10.2196/18189] [Medline: 32663146]
75. Nicholls SG, Carroll K, Nix HP, Li F, Hey SP, Mitchell SL, et al. Ethical considerations within pragmatic randomized controlled trials in dementia: results from a literature survey. Alzheimers Dement (N Y) 2022 May 02;8(1):e12287 [FREE Full text] [doi: 10.1002/trc2.12287] [Medline: 35509502]
76. Benson C, Friz A, Mullen S, Block L, Gilmore-Bykovskyi A. Ethical and methodological considerations for evaluating participant views on Alzheimer's and dementia research. J Empir Res Hum Res Ethics 2021;16(1-2):88-104 [FREE Full text] [doi: 10.1177/1556264620974898] [Medline: 33238781]
77. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1(Suppl 1):S48-S61 [FREE Full text] [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
78. Gordon WJ, Patel V, Thornhill W, Bates DW, Landman A. Characteristics of patients using patient-facing application programming interface technology at a US health care system. JAMA Netw Open 2020 Oct 01;3(10):e2022408 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.22408] [Medline: 33107918]
79. Pais S, Parry D, Huang Y. Suitability of fast healthcare interoperability resources (FHIR) for wellness data. In: Proceedings of the 50th Hawaii International Conference on System Sciences. 2017 Presented at: HICSS 2017; January 4-7, 2017; Hilton Waikoloa Village, HI. [doi: 10.24251/hicss.2017.423]
80. Hampel H, Au R, Mattke S, van der Flier WM, Aisen P, Apostolova L, et al. Designing the next-generation clinical care pathway for Alzheimer's disease. Nat Aging 2022 Aug;2(8):692-703 [FREE Full text] [doi: 10.1038/s43587-022-00269-x] [Medline: 37118137]
81. Saho K, Uemura K, Sugano K, Matsumoto M. Using micro-doppler radar to measure gait features associated with cognitive functions in elderly adults. IEEE Access 2019 Feb 20;7:24122-24131. [doi: 10.1109/access.2019.2900303]

82. Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell 2022;15(1):1-22 [FREE Full text] [doi: 10.1007/s12065-020-00540-3] [Medline: 33425040]
83. Lua EK, Crowcroft J, Pias M, Sharma R, Lim S. A survey and comparison of peer-to-peer overlay network schemes. IEEE Commun Surv Tutor 2005 Jun 30;7(2):72-93 [FREE Full text] [doi: 10.1109/comst.2005.1610546]
84. Suciu G, Stefanescu S, Beceanu C, Ceaparu M. WebRTC role in real-time communication and video conferencing. In: Proceedings of the Global Internet of Things Summit. 2020 Presented at: GIoTS 2020; June 3, 2020; Dublin, Ireland. [doi: 10.1109/giots49054.2020.9119656]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementias
AI: artificial intelligence
BPSD: behavioral and psychological symptoms of dementia
EHR: electronic health record
GDPR: General Data Protection Regulation
HCP: health care professional
HIPAA: Health Insurance Portability and Accountability Act
IoMT: Internet of Medical Things
IoT: Internet of Things
MCI: mild cognitive impairment
mHealth: mobile health
P2P: peer to peer
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RQ: research question
WebRTC: Web Real-Time Communication

Edited by H LaMonica; submitted 23.11.24; peer-reviewed by P Wu, D Viderman, M OldeRikkert, S Groeneveld; comments to author 13.02.25; revised version received 24.02.25; accepted 25.02.25; published 14.05.25.

Please cite as:
Shaik MA, Anik FI, Hasan MM, Chakravarty S, Ramos MD, Rahman MA, Ahamed SI, Sakib N
Advancing Remote Monitoring for Patients With Alzheimer Disease and Related Dementias: Systematic Review
JMIR Aging 2025;8:e69175
URL: https://aging.jmir.org/2025/1/e69175
doi:10.2196/69175
PMID:

Review

Enhancing Enrollment and Adherence in Long-Term Wearable Research on Dementia: Qualitative Systematic Review and Meta-Synthesis

Colleen M Peterson[1], PhD; Renée M St Louis[1], PhD; Carol Flannagan[1], PhD

University of Michigan Transportation Research Institute, Ann Arbor, MI, United States

Corresponding Author:
Colleen M Peterson, PhD
University of Michigan Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI, 48109
United States
Phone: 1 734 489 1709
Email: cmpete@umich.edu

Abstract

Background: With the rapid expansion of wearable technologies, there is increased interest in their utility for passive data collection applications in research on aging. Wearables can be beneficial for research with people with dementia and their families, who have burdens that can make both study participation and reliable data collection more difficult, especially as dementia progresses, but their use also has challenges. Population-specific issues affecting the success of wearables for data collection can include remembering to wear a device, fluctuating acceptance of the device or study participation, and reliance on already strained caregivers.

Objective: This study aimed to systematically evaluate contemporary wearables research to describe persons with dementia's experiences with wearables, their desired qualities, and protocol needs to enhance participant buy-in and sustained wearing for better quality dementia research.

Methods: We used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist for systematic reviews and searched 3 scholarly databases using Medical Subject Headings (MeSH) terms for papers published since 2018 featuring the use or discussion of wearable devices for persons with dementia. We screened 1757 abstracts and retained 58 for full-text review.

Results: We present synthesized preferences, barriers, and facilitators to buy-in and adherence to wearables in dementia research. A total of 29 factors were categorized into 4 overarching categories aligned with study development: device selection, protocol considerations, enhancing recruitment, and promoting adherence.

Conclusions: These findings inform researcher guidelines for wearable device selection and protocol design to enhance the utility of wearables in future longitudinal research featuring persons with dementia and their caregivers.

(JMIR Aging 2025;8:e63768) doi:10.2196/63768

KEYWORDS

dementia; dementia care; Alzheimer disease; caregivers; cognitive impairment; wearable electronic devices; systematic review

Introduction

Background

The growing prevalence of dementia and its substantial economic and social impacts have led to increased attention and funding allocations for research supporting persons living with dementia and their caregivers as well as dementia prevention research [1]. Obtaining the perspective of persons living with dementia and their caregivers is critical to advancing our understanding of dementia to not only find effective treatments but also to address burdens often associated with providing care to this population [2,3]. Engaging persons living with dementia and their caregivers in research studies can be challenging for many reasons, including the fluctuating cognitive capacities of individuals with dementia, the logistical and emotional demands placed on caregivers, and the need for research protocols that accommodate the specific needs and

limitations of this population [3]. Innovative methods to increase meaningful engagement in research studies involving persons living with dementia and their caregivers are warranted.

Meeting Researcher Needs

Wearable technology is a noninvasive, passive way to collect a variety of data over long periods and offers the potential benefit of enhancing care and prevention efforts [4-6]. Remote data collection, such as that offered by wearable technology, has many potential benefits, including reducing recall bias and decreasing the burden placed on participants owing to frequent contact for research purposes. Identifying wearables and research protocols that meet both researcher and participant needs will help wearable technology be more effectively used in longitudinal dementia studies [7]. As the wearable technology market is expected to continue growing substantially [8], guidance for researchers to optimize the selection of devices to enhance buy-in and adherence in their studies is necessary. Wearable devices and research strategies tailored to persons living with dementia and their care providers are crucial for optimizing long-term research with this population [9].

Currently, longitudinal dementia research with wearables faces 3 overarching challenges. First, the challenges particular to persons with dementia and their caregivers and health care professionals must be considered to gain long-term buy-in and use, especially as dementia progresses and caregiving burdens increase [10,11]. Population-specific issues include remembering to wear or charge the device, ethical informed consent [12], and privacy concerns [13,14] that may differ from cognitively intact older adults [15,16]. Second, there is a multitude of wearables on the market with a variety of features and hardware forms at various costs that have not been evaluated in terms of this population's usability and adherence factors. Third, wearables must also provide consistent, accurate, and useful data that meet researcher needs (eg, activity monitors, GPS, and location information) and have limited or easily resolvable technical problems, user-related issues, or data access factors [17].

To date, there has not been a systematic evaluation of wearables for research that holistically examines whether wearables meet the needs of persons with dementia, their caregivers, and researchers to enhance participant buy-in and research utility. Moreover, with the rapid expansion and advancement of wearable technology in the last few years, new research on participant experiences and desires for wearables is necessary. Recent wearables research with this population has featured focus group discussions with limited representation of persons with dementia and their caregivers [10,18], as compared to the experiences of older adults without dementia [19]. Observational research that could inform wearable acceptance and adherence strategies has often been limited to very short-term evaluation periods (typically up to 7 days) with a handful of activity monitoring devices [20-22]. Identifying factors related to wearables that are acceptable to persons living with dementia and their care providers, as well as population-specific support systems, is necessary to strengthen the quality of research requiring long-term adherence [7].

Objectives

This study aimed to conduct a systematic review of dementia-related preferences, barriers, and facilitators to using wearables in research studies to inform future study development, including device selection and participant support protocols.

Methods

Data Sources

This systematic review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist (Multimedia Appendix 1) [23]. Original research studies were identified from the PubMed, MEDLINE, Scopus, and CINAHL databases (all last consulted in January 2023) and Google Scholar. In March 2024, approximately 500 results were reviewed to capture additional references since the 2023 search. The search strategy included a combination of Medical Subject Headings (MeSH) terms and keywords. Categorized according to the PICO (population, intervention, comparison, and outcomes) framework, search terms were used in OR statements in combination for the population ("dementia," "Alzheimer* Disease," "memory loss," and "cognitive impairment") and the intervention or indicator ("wearable*," "wearable sensor*," "tracker*," "biosensor*," "Fitbit," "wearable electronic*," "remote monitoring," and "pedometer"). Regarding comparison, the inclusion criteria did not require intervention trials, whereas for outcomes, the inclusion criteria did not specify the outcomes of the studies.

The full text of the relevant articles was retrieved from the Web of Science, Google Scholar, and library access at the University of Michigan. The articles were reviewed for additional related references.

Study Selection

The selection criteria for the literature search included (1) full-text data from primary peer-reviewed journal articles, commentaries, and conference proceedings; (2) use of remote monitoring via electronic devices that captured data of persons; (3) enrollment of, or reference to, populations living with Alzheimer disease, dementias, or chronic memory impairment or loss impairment; (4) publication in English from January 2018 to March 2024; and (5) documentation of any preferences, barriers, or facilitators to wearables device use. Notably, research featuring Huntington or Parkinson disease was included if participants were experiencing dementia-type impairment and otherwise met the inclusion criteria. Studies were excluded if they (1) did not meet the inclusion criteria, (2) used an implantable wearable or a device not intended for data collection (eg, hearing aids), (3) reported only a study protocol, or (4) included no person interaction (eg, device development without consulting with the population considered affected). Dissertations, books, book chapters, other reviews, and single-case studies were excluded.

CMP searched independently, and CMP and CF reviewed titles and abstracts for relevance and subsequently conducted the full review. CMP manually searched the reference sections of these full-text articles and key reviews for additional citations. RMSL

independently reviewed the search criteria and list of full-text articles to meet the eligibility criteria. The authors used EndNote (Clarivate), a bibliographic tool that allowed them to organize PDFs of articles by topic, along with other data about each article.

Data Extraction and Analysis

Data extracted (by CMP) from the full text of the eligible studies included, when available, publication details, study location, study design, study population, device type and location worn, and device-wearing duration. For the focus of this study, we extracted information related to the preferences, barriers, and facilitators of the use, adoption, maintenance, and adherence of wearable electronic devices by persons with dementia, including ideal device qualities and protocol needs for persons living with dementia and their caregivers. This information could come from the researchers (eg, procedures and processes attempted), persons with dementia (eg, reported preferences and difficulties inhibiting ongoing use), or persons caring for those living with dementia, whether health professionals or informal family caregivers (eg, encouraging persons to wear the device).

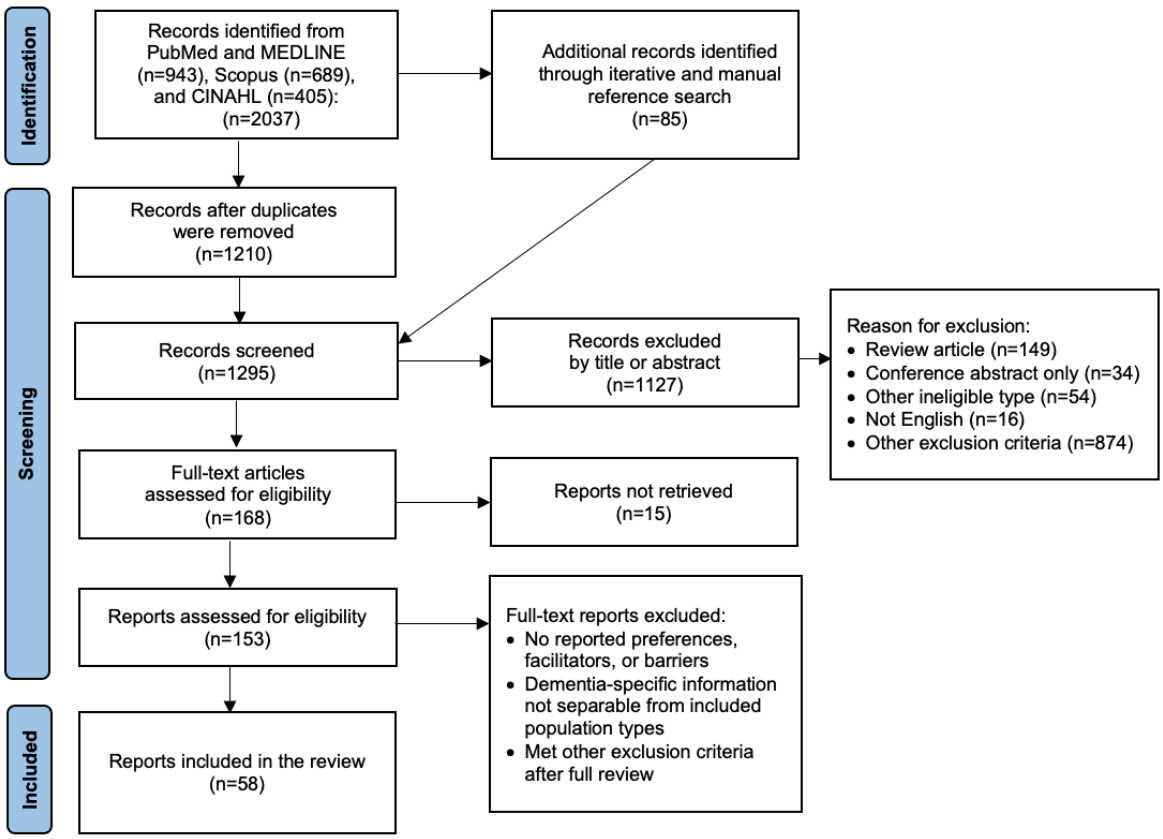
The extracted qualitative data on preferences, barriers, and facilitators were reviewed and initially synthesized for each study by CMP into a common factor coding structure using a grounded theory qualitative research approach [24]. This approach does not presuppose findings but relies on iterative "constant" comparison of collected data to develop codes and overarching themes. RMSL reviewed the summaries and factor synthesis, which CMP and RMSL then independently assigned to each study. The factor assignments were reviewed, and discrepancies were adjudicated together. We report the descriptive frequencies of the factors identified and noted in at least 5 studies and provide a qualitative synthesis of their implications for optimizing long-term dementia research.

Results

Overview

This systematic review identified 1757 studies through the databases as well as iterative and manual reference searches. After the eligibility criteria were applied, 3.3% (58/1757) of the studies were retained for data extraction analysis (Figure 1 [23]). Most studies (1127/1295, 87.03%) were excluded based on the abstract review because the device evaluated was not a wearable or the population had some cognitive impairment but did not specifically include dementia. The studies were primarily observational (eg, feasibility studies) and qualitative (eg, using interviews or focus groups), with only 19% (11/58) featuring some type of trial. Participants were most often persons living with dementia in the community (as opposed to residential care), their informal family care partners, and paid dementia care health professionals. The duration for which persons living with dementia were asked to wear the study device ranged from a few hours over several days to 12 months, though 7 days was a common wear period. Information pertaining to the individual studies, including the type of wearable, is summarized in Multimedia Appendix 2 [9,10,18,20,22,25-60].

Figure 1. PRISMA 2020 flow diagram for the systematic review of wearables in dementia research.



Barriers, facilitators, and preferences from the 58 studies were qualitatively synthesized into a total of 29 factors that fit into 4 overarching themes. These themes are categorized by issues dementia researchers should consider from study development to deployment: device selection, protocol considerations, enhancing recruitment, and promoting adherence. Although we categorized these factors by greatest relevance to a particular study section, it should be recognized that many of the factors are interconnected and related. The frequency and proportion of studies featuring each factor are detailed in Table 1. The factors present in each study and their cross-reference numbers are available in Multimedia Appendix 3 [9,10,18,20,22,25-77]. As we describe the factors in the results, we use the term caregivers inclusively to mean both informal (ie, unpaid, often family members) and paid caregivers (eg, health professionals and residential care staff) and provide relevant quotes that capture participant experiences. In the Discussion section, we offer a synthesis of recommendations for researchers planning to conduct longitudinal research with populations with dementia.

Table 1. Presence of synthesized factors across the studies (N=58).

Factor	Count, n (%)
Device selection	
Easy to use	21 (36)
Smaller size or weight	18 (31)
Comfort	15 (26)
Fits into routine	13 (22)
Unobtrusive	12 (21)
Tailorable	12 (21)
Materials	10 (17)
Aesthetics	8 (14)
Easy to wear	7 (12)
Water resistant	7 (12)
Protocol considerations	
Privacy concerns	15 (26)
Consent capacity	14 (24)
Battery	13 (22)
Adjustment period	10 (17)
Task requirements	9 (16)
Enhancing recruitment	
Technical anxiety	17 (29)
Multifunctional	16 (28)
Burden impact	16 (28)
Stigma	13 (22)
Caregiver buy-in	12 (21)
Connectivity	10 (17)
Promoting adherence	
Remote monitoring	19 (33)
Provide health insight	18 (31)
Technical support	17 (29)
Caregiver support	14 (24)
Safety	14 (24)
Remembering the device	12 (21)
Independence	10 (17)
Self-removal	7 (12)

Device Selection

Easy to Use

Given the care burden of family members or paid caregivers in residential settings, a wearable's ease of use was paramount when asking caregivers to be involved in a study. It was the most frequently identified factor in studies, with 36% (21/58) of studies mentioning it. The wearable device must be easy to use [25-27,61-65] with an intuitive interface and features [28-31]. The device should require little maintenance or technical support for everyday use, and any interaction with the device required by the study should be simple [10,28,32-35,66]. Some studies explicitly noted that perceived ease of use was associated with more positive attitudes toward wearing the device and less effort or sense of inconvenience in participation [27,36-39]. As a person living with dementia exclaimed in the study by Gris et al [38], "Older people need to use simple things!"

Smaller Size and Weight

Almost without exception, participants desired smaller and lighter wearables or expressed discontent about the size and weight of the device they were asked to wear in the study [34,35,40-42,63,64,66-68]. The influence of a wearable's size and weight was the most frequent physical form factor, with 31% (18/58) of the studies including comments about it, and it was the fourth most common factor identified overall. While rare, there were some exceptions to participants desiring smaller devices overall. Some suggested enlarging certain facets of a device, such as having a larger face or font to read [43] or larger buttons to interact with [27].

Comfort

The physical ease of wearing the device was important [28,31-33,38,41,44,67,69], with discomfort specifically cited as a reason why participants removed the device in several studies [34,45,66]. *Comfort* was sometimes explicitly linked with *unobtrusive*, as complementary qualities [20,29,61]. Nevertheless, prototype devices that were bulky or wearables with a more invasive application (eg, required taping to limbs or backs) were mostly tolerated in the short term. However, simply put, "If residents would show discomfort related to the wearable, the wearable would not be used" [61].

Fits Into Routine

Studies found that if a device type was already familiar or similar to what caregivers and persons living with dementia already used regularly, the wearable could be seen as convenient and more easily fit into a participant's lifestyle [10,30,32,33,46,67,70]. Wristwatches were the most frequently cited as meeting these criteria [20,29,35,38,41,47]. At the same time, devices might also be removed if accessory removal is also part of a participant's routine [20,32,41]: "To tackle such barriers [to using mHealth devices] will likewise require (informal) caregivers to change their attitudes towards using digital technologies and the integration of these technologies into routine care" [30].

Unobtrusive

A wearable that is not obvious or can be hidden was considered a distinct positive [25,29,35,38,70,71]. Having the ability to conceal a wearable alleviated some worry about a person living with dementia being resistant to wearing the device or removing it (refer to the Self-Removal subsection), which could be especially pertinent in studies involving individuals with more advanced stages of dementia [61,66,68]. Unobtrusive devices were also seen as reducing the potential of drawing attention to themselves or experiencing stigma associated with wearing a device [10,42,61] (refer to the Stigma subsection). Like the *fits into routine* factor, wristwatches were commonly cited as an unobtrusive solution [10,20], but something that could be slipped onto a belt or into a shoe was also noted as a viable option [10,29].

Tailorable

Having the option to use a wearable in a way that fits a person's particular style preferences was desired, from format (eg, neck pendant or watch) to style (eg, colors and design) [10,29,34,35,38,61,62,71], as some had concerns with the particular format that was provided in the study and desired a different way to wear the device [26,31,68]. Alternative device forms can include a spectrum of participants and their day-to-day presentations (eg, gender and culture) [10,18]. Offering a wearable in several different forms could help a device suit the *fits into routine* factor, but the *tailorable* factor highlights that singular forms may not fit into everyone's routine equally: "Several members in all four groups expressed a desire for more options in device styles...all groups asked if alternative styles could include a broach, belt clip, small bracelet, or ankle bracelet" [31].

Materials

Nonspecific material concerns or durability were cited as important considerations for wearables to be satisfactory to participants and researchers alike [29,41,63]. Specific material issues such as softness [34,63,66] and the need to make wearables safe for those with skin sensitivities or allergies were also mentioned [10,28,34,61]. Some participants stopped wearing the device altogether because it caused irritation [22] or even skin abrasions [48].

Aesthetics

The general appearance of the device design in its look and feel could be a barrier to everyday use and, thus, adherence [41,43,47,63,64]. Size, form, color, and device material all play a role in the aesthetic attractiveness of a wearable [31,34,35]. For example, participants in the study conducted by Jacklin et al [31] were dissatisfied with the prototype wearable they were testing, describing it as "big and clumsy looking" and "rubbery."

Easy to Wear

Diminishing fine motor skills in persons living with dementia made *easy to wear*, or how simply one can put on or take off the wearable, an issue for day-to-day use [10,62,65]. Wearables that were difficult for the wearer to put on or take off by themselves required caregiver support that could add stress to both the person living with dementia and their caregiver,

creating participation burden and a sense of lack of independence [34,48]. Some wearables were slightly more complicated and required specific sensor placement, which was either difficult to apply correctly or failed to remain securely attached [34,49,50].

The factor *easy to wear* is distinct from *easy to use* and may not play a role if a wearable is designed to be worn continuously. For example, a family caregiver reported that an accelerometer was "a little bit difficult to put on, so [the persons living with dementia] did do it but...it's not easy." The researchers echoed that "the wristband of the accelerometers could be difficult to fasten" but reported that afterward, "the accelerometers were not burdensome to wear" [65].

Water Resistant

Water-resistant or waterproof wearables were recommended [28,61,64] to minimize potential disruptions to data collection. Along with charging the device, a common point of noncompliance occurred when participants forgot to put the wearable back on after removing it for bathing (refer to the Remembering the Device subsection) [29,34]. By selecting a waterproof device and asking participants not to remove it, study protocols do not have to rely on persons living with dementia or their caregivers remembering to take off the device during bathing and, therefore, can reduce the number of study *task requirements* and related *burden impact* [44,63]. Cruz-Sandoval et al [29] suggested that minimizing the need to remove a device can reduce instances of "resistance" from persons living with dementia, and therefore, waterproof devices "can ease some of the burden on caregivers" as well.

Protocol Considerations

Privacy Concerns

Many studies noted that privacy was an issue, with participants voicing concerns about data security and who had access to the wearable's data [10,18,26,31,38,63,68,71]. For example, a caregiver's spouse worried about third-party breaches of her wearable data, explaining that "privacy and security are important to his wife because '...she understands just how easily those systems can be hacked'" [71].

At the same time, some persons living with dementia and their caregivers stated explicitly that privacy concerns were outweighed by the personal or societal benefits of their study participation [25,33], or they reported no privacy concerns [27]. Participants understood that data sharing was often necessary to meet the goals of the study, and privacy concerns were ameliorated when researchers established trust around data management: "Both informal caregivers and the nursing staff were not concerned about privacy issues regarding the wearable, as they consider biomedical variables as de-individualized values. The participants trusted the researchers and nursing home regarding good data handling, as they felt well-informed" [61].

Researchers were able to build trust and attenuate privacy concerns through transparency about data access and security measures during the informed consent processes and throughout the study [61,72]. Empowering participants to have control over privacy options (eg, allowing the collection of some but not all data) and offering the ability to turn off or select what data to monitor also helped [51,73].

Consent Capacity

The ability of persons living with dementia to understand and actively affirm their study participation is an ethical imperative but was specifically mentioned as a potential difficulty by several studies [38,71]. Especially in longer-term studies, seeking ongoing consent to participate was highlighted as necessary, with suggestions to check in regularly to assess the cognitive capacity of the persons living with dementia and their willingness to participate [20,51,74]: "Fluctuating cognition can be a significant barrier to the remote monitoring of isolated patients as forgetting about one's motivations to be monitored can lead to anxiety and agitation for being watched and, consequently, the disablement or destruction of equipment" [74].

At odds with fully informed consent was the general awareness of persons living with dementia regarding their participation or the device [32,43]. At this level of impairment, consent from a caregiver or another proxy for the persons living with dementia was sometimes sought instead, which could be an appropriate barrier to inclusion [22,69,70]. Another major challenge is reconciling what to do when a person living with dementia desires conflict with the caregiver's desire for them to engage in the study and wear the device [10,68]. Under consideration is the issue of dignity, respect for autonomy, and how ignoring these or not obtaining assent could lead to agitation: "In one case, [the study staff] visited the participant after he had refused to wear it on his wrist, saying his wife was tagging him 'like a dog'" [68]. In another study [10], a caregiver reported "[My husband] would have never given his consent even at the beginning—he liked the independence of making his own decisions and doing what he wanted. And I would have had to do it without his consent."

It was recommended that, when possible, researchers, caregivers of persons living with dementia, and their proxies should attend to their care recipient's wishes, whether from an explicit advanced directive [73] or knowledge of their comfort with similar technology [30].

Battery

The longevity of battery life and the ease of charging were noted as concerns for both researchers and participants [9,28,38,63,64,75]. Wearables with a battery life of at least several days were ideal [28,29,52,63], decreasing the need for participants to remember to charge the device [9,10,72] (ie, simplifying *task requirements*), which also minimized data collection interruptions for researchers. Low battery warnings from the device or study staff were seen as a useful tool to ensure an adequate charge of the battery [10,32,68]. Necessarily, the burden of charging was exacerbated when a caregiver was not on site to physically assist or when the charging mechanism was difficult to implement [35].

Adjustment Period

Participants may need time to acclimate to wearing the device (eg, establish wearing routines) and to resolve technical issues before consistent adherence and, therefore, consistent data quality, as adherence tended to improve after an initial period [9,29,32,33,41,53,61]. On the other hand, some studies found better compliance earlier in a study period [42,54,65], perhaps because of novelty. Adjustment periods featuring cognitive aids such as instructional pamphlets and study staff check-ins provided early troubleshooting and time to build new routines [9,29,33].

Task Requirements

Cognitive impairment and caregiving burden were barriers to protocol adherence when participants were asked to do more than simply wear the device [55,56,74,75]. When a study protocol required additional activity from either the persons living with dementia or a caregiver, such as completing bedtime and wake-up event markers, they were less likely to remember or be able to comply [9,50,57,72].

Enhancing Recruitment

Technical Anxiety

Although many interacted with the study wearables and related technology platforms successfully, participants' sense of a lack of technical proficiency and negative attitudes about technology were reported as a perceived barrier to enrollment [10,18,27,37,38,41,71] and adherence during the study [29,53,62]. Caregivers and persons living with dementia alike expressed concern about the complexity or technical know-how needed to operate the wearables [29,36,53,63,74]. Technical anxiety could present as worrying about damaging or losing the device [20,29,34]. The level of technical anxiety and attitudes toward the technology were related to age qualitatively and statistically [10,27,38,53,62] and previous experience with devices [36,46,76]. Participants in a study actually rated themselves as less technologically proficient at the end of the study because of difficulty with the wearable [42].

Multifunctional

Participants saw the potential for the device and platform to offer more information or other uses to the wearer as a benefit of being involved in the research [27,71]. Being able to track their health and activity, signal health events or emergencies (eg, wandering or fall alerts), provide directions, or offer location services, among others, were cited as desirable features [10,25,27,29-31,34,63,64,67,74]. Simpler functions such as information about time, date, and alarm capacity were also ideal wearable facets [10,32,35,41,52,63]: "Participants with dementia indicated that multiple purposes for the locator device, such as being able to measure heart rate, blood pressure and activity level, is advantageous as it would streamline the number of devices that they and their caregivers must contend with" [10]. Stavropoulos et al [63] reported, "Feedback suggested that some people with dementia may be more willing to accept technology that supports them in their daily functioning, in addition to assessing it."

Burden Impact

Caregivers especially liked the potential the wearables offered in terms of supporting their caregiving role and tasks, which, in turn, reduced related burden and stress [27,32,39,64,69]. Factors such as a wearable's capacity for *remote monitoring* and *providing health insights* could improve quality of life or provide peace of mind [33,39,63,67,72]. However, when the wearable was technically difficult to use (*easy to use)* or researchers required additional time or study activity (refer to the Task Requirements subsection), being involved in the study was perceived as a cause of additional burden and a deterrent to both enrollment and continued participation [29,32,33,58,59,61,70,74]. Notably, participants in several studies reported experiencing a mix of both reduced and heightened burden, which varied according to specific facets of the research or from using the wearable [32,33,65,69]: "The nursing staff indicated that they did not want to spend valuable care-time if the foreseen benefit of the wearable for the specific resident was not evident to them. Likewise, also informal caregivers found it important to know the potential individual benefits, before deciding whether to bother the resident with the wearable, as they anticipated some level of resistance to wearing the wristband by the residents" [61].

Stigma

Caregivers and persons living with dementia expressed concerns about the stigma associated with the wearable and the wearable identifying their care recipient (or themselves) as affected by the hidden disability [10,18,30,64]. There was concern about social acceptance and the visibility of the device, as well as the perceptions or comments that it might engender (eg, that they are under surveillance or need to be monitored) [33,34,61,72]. On the other hand, "social influence" was often a positive, with care providers or fellow persons living with dementia offering positive comments or encouraging the use of the wearable [29,37,39,76].

Families and professional care providers alike were sensitive to the negative connotation that some wearables might cause. An Alzheimer Society staff member reported that some families might avoid incorporating a MedicAlert bracelet in their care management plans because "It's a signifier or something that's perceived to bring stigma—[the persons living with dementia] don't want to be out there clearly marked in some way, that there's something wrong with them. Nor does the family" [18].

Caregiver Buy-In

Caregiver buy-in underlies many facets of wearable studies featuring persons living with dementia, as many studies relied on caregiver commitment and coordination with the researchers to do study activities (refer to the Caregiver Support subsection) [20,36,57,61,70]. Moreover, they played significant roles in facilitating understanding of the study by persons living with dementia [29,47] and in encouraging them to wear the device [39,41]. Some studies specifically cited the need for caregivers to have perceived usefulness or reasons to commit to reducing perceived potential burden or risks to enroll themselves and their care recipient in the study [61,72,76]. Gaining caregiver buy-in could be enhanced by reducing participation burdens

(eg, fewer *task requirements*, *easy-to-use* devices, and offering ongoing *technical support*) and experiencing personal benefit from the study (eg, *remote monitoring* and *providing health insights*).

Connectivity

Passive, remote monitoring is the defining advantage of wearable device studies. However, technical connectivity requirements can be a barrier to enrollment [10,31,61,69,74] and may create logistical problems for data transmission [18,31,56,77]. For example, requiring faster Wi-Fi or consistent Bluetooth linkage with a smartphone for more frequent data uploads [67,75,77] may introduce recruitment bias to individuals of higher socioeconomic status and those who are already technically inclined: "The lack of internet connection in patient homes and of funding for caregivers and family member to purchase assistive or communication devices are frequently coupled with skepticism or low abilities to engage with digital products" [74].

Promoting Adherence

Remote Monitoring

Remote monitoring was the most frequently cited desirable and useful feature (mentioned in 19/58, 33% of studies) when engaging in a wearable study among populations with dementia [10,25,27,34,59,61,69,76]. Geolocation-based remote monitoring enabled caregivers to track the movements of persons living with dementia when they were not present, including movement outside the home. Remote monitoring was also seen as a way to know about the well-being of persons living with dementia, for example, by the identification of agitated movement or interrupted sleep [32,40,77]. Remote monitoring enhanced the *independence* of the person living with dementia [27,38,63] while providing informal caregivers with peace of mind regarding the safety of persons living with dementia (eg, reduces *burden impact*) [46,63,64,67,72].

For some participants, the utility of the benefits of geolocation services conclusively outweighed potential *privacy concerns* [25,42]. However, for some, introducing the idea of being monitored could be seen as an infringement on autonomy and privacy [25,31], such as when "some participants feared that proud Anishinaabe older adults would be offended by the idea that their movements would be tracked" [31].

Provide Health Insight

Providing health insight to the participants was one of the top factors mentioned as a benefit to ongoing participation and use of a study's wearable. Paralleling the *multifunctional* factor as an aid to enrollment incentives, health insights included, for example, information on blood pressure, heart rate, sleep quality, and steps. Having ongoing access to this health data motivated participants' personal activity, helped caregivers monitor for health crises, and provided more general insight into the well-being of persons living with dementia [10,25,27,29,32,33,38,53,54,58,61,63,69,74,77]. In addition, they appreciated that this health information could be useful to residential staff and physicians in informing care decisions [25,27,32,34,61,64,69]. A caregiver in the study by Lazarou et al [25] said that with the information from the wearable, "we could monitor her daily activity and communicate any problems to the doctors. It helped us a lot to understand what is going on."

Technical Support

As *technical anxiety* was a barrier, providing initial technical training and ongoing technical support was identified as a complementary facilitator [20,26,37-39,50,74]. Staff availability to provide training at enrollment, support during an adjustment period, and assistance as needed throughout the study helped improve adherence [29,33,77]. Stand-alone instructional materials and cognitive aids were useful to participants who sought to troubleshoot issues on their own [9,10,47,66]. Along with data monitoring, regular proactive check-ins with participants made (or would have made) staff aware of technical problems and helped maintain better compliance [29,45,53,68,77].

For example, Gelonch et al [33] provided caregivers with initial training and ongoing technical assistance, and consequently, they supported their care recipient's use of a lifelogging camera: "The [persons living with dementia] said that the support and supervision provided by their caregiver in certain tasks or actions, such as connecting and loading the camera, were also key for them continuing with the experiment."

Even when apprehensive about using the technology (refer to the Technical Anxiety subsection), participants commonly showed a readiness to learn how to operate the wearable devices and follow study protocol. As recounted by Anderson et al [36], "Many were not necessarily technologically proficient caregivers, but they wanted to be useful. None of the caregivers implied that the research should not continue to be refined."

Caregiver Support

While the importance of *caregiver buy-in* was highlighted in recruitment, in this section, we underscore the importance of caregivers and their investment in the study for continued wearables adherence [36,39,71,77]. Many research protocols required caregiver activity to accomplish their goals by having them be the primary study participant or helping persons living with dementia complete study tasks [29,44] and, importantly, charging the wearable [35,77]. Many factors depend on the caregiver, such as helping persons living with dementia *remember the device* [20,41], assisting with putting it on (refer to the Easy to Wear subsection) [34,48], or understanding how to operate the wearable [33]. In addition, caregivers could be instrumental in explaining facets of the study and, in other ways, facilitating communication of research needs to their care recipient throughout the study [29,68].

Notably, relying on caregiver support can be a major ask (ie, increase *burden impact*): "The caregivers were able to be present almost all the time. They demonstrated personal investment in the care that the person living with dementia received. Their willingness to participate in, work with researchers, and use technology 24 hours a day for 30-60 days was a significant commitment" [36].

Safety

Broadly speaking, wearable technologies were seen as an instrument for enhancing the safety of persons living with dementia [45,46,71,72]. This potential positive was specifically identified as an ongoing benefit, often in tandem with or implied within *remote monitoring* and *providing health insight*. Wearables offer a sense of security to persons living with dementia and their care providers, aiding aging in place by alerting caregivers to life events or potential crises (eg, elopement or fall monitor) and detecting changes in health or behavioral patterns [28,31,34,59].

On rare occasions, the device itself was harmful, such as when an over-ear device caused skin abrasions [48] or when participants desired a different product or sensor design for safety reasons [69]. In contrast to other participants in their study, a caregiver in the study by Snyder et al [71] explained that they felt the wearable, an emergency response pendant, gave a false sense of security, saying, "My reluctance is that it doesn't seem...It can't stop something from happening to her. It just alerts you maybe if something does."

Remembering the Device

Adherence in wearable studies was affected by persons living with dementia's ongoing ability to remember what the device was for [22,29,42]. Not remembering the purpose of the device could also result in "confusion and agitation" [43]. More commonly, persons living with dementia simply forgot to wear the device [10,20,63,75], especially after removing it for bathing (if not *water resistant*) or sleep, often resulting in failure to put it back on [34,65]. Better adherence was reported earlier in the day [53,77], potentially because of sundowning: "Due to memory impairment participants had to be constantly reminded to wear the device and they could change their opinion from one day to the other. Activities that made wearing the device more salient, such as the need to take it off to charge it or bathing, proved challenging with some participants" [20].

Independence

Whether the device provided geolocating, physical activity monitoring, or health alerts, participants, especially caregivers, saw these devices and platform features as enhancing the independence and autonomy of persons living with dementia [10,27,39,63,67,71]. Participants' ongoing access to the data and ability to turn off data collection at their discretion contributed to a sense of agency within the study itself [33,38]. Similar to the potential harms of *remote monitoring*, Synder et al [71] framed this positive as part of the "conundrum between the benefits of having peace of mind, a sense of self-efficacy, and greater independence versus perceiving (remote monitoring technologies) as an invasion of privacy, a security risk, a false sense of security, and a distraction."

Self-Removal

Taking off the device when not instructed was an issue in several studies [44,60]. This could be temporary, such as during bathing, or routine, such as removing accessories before sleep [32,72]. In some cases, removal was longer term, often due to agitation (possibly related to sundowning) or discomfort [22,32,66]. Some researchers and care providers suggested or chose to put the wearable out of reach (eg, on their back) to prevent self-removal by persons living with dementia [60,61].

Discussion

Principal Findings

Overview

This review identified barriers, facilitators, and preferences that could improve the quality of research studies featuring wearable devices and persons living with dementia. The review included studies encompassing a variety of study designs and perspectives from researchers, persons living with dementia, and their caregivers (both unpaid family members and paid professionals), ensuring broad inclusion of research experiences. We offer a qualitative synthesis of these findings to enhance enrollment and adherence in long-term wearable research featuring persons living with dementia. Recommendations for future research fall within 3 main categories: device selection, ethical study design, and technical and data management plans.

Device Selection

Dementia Impairment Level Requires Different Devices

We identified the importance of selecting multifunctional devices with the capacity for participants to monitor their own or their care recipient's data, whether on the device itself or via a dedicated platform. However, what persons living with dementia and their caregivers preferred in terms of devices was sometimes different depending on the advancement of dementia in their target population. In earlier stages, persons living with dementia themselves were more engaged and desired more interactive, multifunctional devices that provided personal benefit to them. However, persons living with more severe dementia, from their perspective and their caregivers' perspectives, desired or interacted with only simpler wearable form factor and features (eg, just used the clock feature) [77]. Even less obtrusive devices or ones out of reach may also be warranted if they are more likely to remove the device themselves as dementia progresses.

The growing ubiquity of remote monitoring and wearable devices for everyday use likely makes stigma less of a concern as we move into the future because an older adult will not stand out when wearing something such as that, especially if it is unobtrusive or looks like an everyday wristwatch. The availability of the devices likely also drives expectations of what participants will personally gain from being in a research study, as people are more familiar with what type of information gathering is theoretically possible, and participants may expect to have access to that information too.

Wristwatches Fits Into Routine but Could Also Be Problematic

Wristwatch wearables were highly recommended as they are part of normal wear for many and offer different levels of multifunctionality. In addition, their unobtrusiveness can reduce stigma [10]. However, habitual device removal during sleep and bathing often conflicted with the research protocols requiring participants to wear the devices all the time [32,41]. Waterproof forms, along with clear instructions for the caregiver

and persons living with dementia to avoid removal in these situations, are useful. Researchers should understand that adjustment periods may be warranted with proactive reminders as participants acclimate to the newer routine and have fewer persistent issues [41].

Ethical Study Design

Attend to Fluctuating Consent Capacity in Persons Living With Dementia

Persons living with dementia feel actively engaged in the research and empowered over their health status when they can see their health data and decide when to engage in the study. However, progressive impairment can mean a person living with dementia may experience fluctuating ability to consent and understand the protocol for a long-term study. Researchers in longer-term studies should establish ongoing communication to monitor possible changes in device preferences or needs. Advanced cognitive decline from dementia can mean decision-making and consent ability are totally compromised for the person living with dementia. Reliance on a caregiver or another surrogate for consent is common, but studies emphasized that if the person living with dementia has given their thoughts on similar situations before, those preferences should be prioritized. Respecting participants' agency is paramount; recognizing and responding to the issues of ongoing consent in this population is key to conducting an ethical study. Actively checking in with participants on a regular basis in long-term studies to gain assent or consent for continued data collection is ethical and may also be required by institutional review boards, given expected cognitive changes.

Promote Inclusivity With Wearable and Protocol Choices

Researchers should be aware of how their study protocol and wearable device requirements may impact who can or would enroll in their study. Some studies, such as that by Cohen et al [53], suggested screening participants for technical knowledge and ability as a way of ensuring compliant participants are enrolled. However, this is not inclusive and could result in bias in recruitment. Instead, preexisting technology and connectivity needs (eg, stable Wi-Fi) should be minimized or eliminated to enhance distributive justice [73,78] and increase the representativeness of the study results. Researchers can reduce the burden and increase equity by not requiring constant wireless connectivity or by supplying the technology to participants. However, intermittent data collection may be at odds with proactive data monitoring for troubleshooting purposes. Staff phone calls or other lower-technology check-ins are recommended in this case.

If possible, different formats and styles (eg, watch vs pendant and watchband options) that are familiar and fit a participant's personal, cultural, and gender preferences should be offered [79]. However, it is important to balance these preferences with the type and quality of data that each format can reliably collect [34]. As evidenced by this review and other research efforts, user-centered codevelopment of wearables with persons living with dementia for research purposes is feasible [38,80] and should continue to be explored to enhance recruitment and adherence in longitudinal studies featuring persons living with dementia.

Data Management

Implement Data Management Plans

We categorized a total of 29 factors into 4 categories that aligned with study development (device selection, protocol considerations, enhancing recruitment, and promoting adherence). However, a fifth category—data management—was largely left uncommented on, as studies were primarily focused on the needs and preferences of the participants rather than the researchers. Participants may report ongoing adherence, but making sure a wearable is not only worn but also collecting data is paramount to conducting quality longitudinal research.

Researchers' data management plans should comprise data collection, storage, and processing for analysis [29,73]. For example, researchers conducting long-term studies should know how long their device batteries will last under their protocol and have procedures in place to limit both data loss and participant burden (eg, reminders to participants to charge their devices). By regularly checking data for completeness, researchers can identify technical problems and offer troubleshooting to resolve missing data issues. Researchers should also have a loss prevention plan for the wearable to reduce missing or inactive devices and the resulting data loss.

For example, Cruz-Sandoval et al [29] had a comprehensive data management plan for a wearable study in an assisted living setting. They combined weekly program visits with battery charging, data collection (ie, upload and synchronization with the study server), and addressed technical troubleshooting needs. They also reviewed the uploaded data for abnormal data patterns and double-checked with caregivers "as soon as possible" to understand if the abnormal data were valid or if they should flag it for deletion. As a loss prevention plan, the researchers also asked the staff to report on the location of the device at the end of each work shift. Particular data management plans could also influence device options; Muurling et al [81] recommended that real-time participant access to device data could both increase adherence and help alert researchers to data issues in close to real time.

Collective Data Management Efforts Could Enhance Support for Wearables Research

To reduce researcher data burden in individual studies, research groups could build infrastructure and codebases to share collectively in an open-access format (eg, on GitHub). Although researcher perspectives on the use of the devices and subsequent data management were generally missing in the studies, 1 study reported that their experience organizing, cleaning, and analyzing Fitbit data were "onerous and time consuming" and recommended that others use automated data management software [77]. Researchers should be aware that many commercial, proprietary wearables may not offer researchers access to the raw data, which may limit secondary utility and increase reliance on the original developers for measure validation. Other collective data enhancements include setting standard definitions or ideal compliance goals, which varied across studies or were not reported at all. Clear, agreed-upon

definitions for adherence and reporting standards are critical for improving long-term dementia research with wearable devices [9,47]. Indeed, a lack of specific, consistent adherence and attrition data from many studies made quantitative comparisons related to the factors impossible. When combined with real-time data monitoring, shared data expertise can help researchers determine whether seemingly incoherent data result from technical issues, allowing for timely resolution [81], or from changes in cognitive function and life circumstances as dementia progresses [82].

Informing Wearable Technology Adoption and Use Models

This systematic review evaluated preferences, barriers, and facilitators to wearables in a study setting, so everyday technology adoption measures may not be entirely applicable. Indeed, caregivers and persons living with dementia may endure adjustment periods, committing themselves to using difficult technologies if the knowledge gained will advance dementia prevention efforts or benefit other caregivers [82]. However, because of the lack of consistent quantitative data, we could not assess factor relationships quantitatively with wearable technology adoption and use models such as the unified theory of acceptance and use of technology model [83], its extended version—the unified theory of acceptance and use of technology 2 model [37], the wearable acceptance in health care model [84], or the technology acceptance model 3 [85]. These measures may also need to be updated as overall perceptions of technology change and its use becomes more ubiquitous, even among older adults. Surveys and other marketing research indicate that although fewer older adults use wearables compared to other age demographics, approximately 25% of US adults aged ≥65 were using wearables in 2023 [86], a rapid increase even in just the last few years [87]. Moreover, these measures relate to technology acceptance in general, while motivations underlying intentions and actual use of wearables may be different, especially as noted within research contexts. A research-specific technology acceptance questionnaire could help researchers understand perceptions directly related to their needs and identify additional barriers to recruitment and adherence.

There is also utility in a more systematic assessment of wearable device qualities themselves and how they relate to adherence. The use of measures such as the Quebec User Evaluation of Satisfaction With Assistive Technology (QUEST; version 2.0) can simplify device comparisons and inform future wearable choices or commercial development. Future reviews could also benefit from standard reporting of why particular devices were chosen in original studies.

Clinical Implications

We have reviewed and identified factors related to the use of wearables in long-term dementia research. However, wearables are also a potentially viable way to gather data for clinical decision-making because they can offer a way for clinicians to proactively monitor health data and circumvent some of the challenges around gathering a patient history from someone with dementia [64,88]. Supporting this potential, a large cohort study of cognitively healthy participants showed that changes in long-term 24-hour activity patterns were associated with elevated risk for the onset of Alzheimer disease, Parkinson disease, and cognitive decline [89]. Moreover, as wearable devices are being used more by the general population, clinicians are experiencing patients coming in with questions about what they have learned from their own wearable devices. Improving research in this area can help clinicians think about how they can interpret and use wearable data in a meaningful way (eg, personalized medicine). Clinicians could use the results of this paper to guide a conversation with their patients about selecting a wearable that is beneficial for meeting the needs of both the patient and their health care team, which could also enhance adherence.

Limitations and Strengths

Our findings are generally agnostic of cognitive impairment status or degree of dementia, reflecting the participant mix seen in the studies, which also often did not disambiguate by the severity or type of dementia (eg, Alzheimer disease and Lewy body disease), which may influence findings as well as the generalizability of findings. Because our analysis was not limited to specific device types or study goals (eg, physical activity or dementia prevention), not all factors may be applicable to all future research applications. We did not perform a formal risk of bias assessment because of the heterogeneity of the study types included, and we note that many studies may be considered at high risk of bias because of smaller sample sizes. However, considering the outcomes of interest for this study were qualitative in nature, the samples were considered sufficient.

Because our extraction captured all reported barriers and facilitators regardless of their origin, we identified the holistic perspective of researchers, persons living with dementia, and their care providers, including unpaid family members or paid health professionals. We systematically coded barriers, facilitators, and perspectives using a grounded, qualitative approach, which did not presuppose expected outcomes, allowing for new or unexpected insights. We offer a comprehensive synthesis of multiple types of studies featuring many wearables and outcomes of interest, identifying overarching themes core to long-term wearable research featuring populations with dementia.

Conclusions

We described factors critical to enhancing dementia study recruitment and wearable adherence, emphasizing how those factors should be considered collectively in an ethical context to meet both researcher needs and participant desires. Organized by study design order and literature prevalence, dementia researchers can use the findings of this systematic review to thoughtfully design quality, long-term wearable studies. The incorporation of wearable technology in dementia research presents a promising pathway to surmount traditional challenges associated with studying persons living with dementia and their caregivers. By leveraging these noninvasive tools, researchers can obtain continuous, objective data for use in advancing treatment, care, and prevention efforts. The results of this evaluation will facilitate quality improvement in ongoing research efforts and contribute to a more person-centered approach to developing effective interventions and prevention strategies.

Acknowledgments

This work was presented in part at the annual meeting of the Gerontological Society of America. This study was supported by the University of Michigan Network for Innovative Methods in Longitudinal Studies of Aging through the National Institute on Aging (grant #1 R24 AG077012-01). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute on Aging.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA 2020 checklist.
[PDF File (Adobe PDF File), 110 KB - aging_v8i1e63768_app1.pdf]

Multimedia Appendix 2
Summary of included wearable research studies involving populations with dementia.
[DOCX File , 38 KB - aging_v8i1e63768_app2.docx]

Multimedia Appendix 3
Presence and frequency of factors across wearable research studies involving populations with dementia.
[DOCX File , 65 KB - aging_v8i1e63768_app3.docx]

References

1. GBD 2016 Dementia Collaborators. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 Jan;18(1):88-106 [FREE Full text] [doi: 10.1016/S1474-4422(18)30403-4] [Medline: 30497964]
2. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
3. Frank L, Jennings LA, Petersen RC, Majid T, Gilmore-Bykovskyi A, Schicker L, et al. Participation of persons with dementia and their caregivers in research. J Am Geriatr Soc 2021 Jul 10;69(7):1784-1792 [FREE Full text] [doi: 10.1111/jgs.17340] [Medline: 34245586]
4. Chen C, Ding S, Wang J. Digital health for aging populations. Nat Med 2023 Jul 18;29(7):1623-1630. [doi: 10.1038/s41591-023-02391-8]
5. Cullen A, Mazhar MK, Smith MD, Lithander FE, Ó Breasail M, Henderson EJ. Wearable and portable GPS solutions for monitoring mobility in dementia: a systematic review. Sensors (Basel) 2022 Apr 27;22(9):3336 [FREE Full text] [doi: 10.3390/s22093336] [Medline: 35591026]
6. Husebo BS, Heintz HL, Berge LI, Owoyemi P, Rahman AT, Vahia IV. Sensing technology to monitor behavioral and psychological symptoms and to assess treatment response in people with dementia. A systematic review. Front Pharmacol 2019 Feb 4;10:1699 [FREE Full text] [doi: 10.3389/fphar.2019.01699] [Medline: 32116687]
7. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023 Jan 8;11:1281923 [FREE Full text] [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
8. Hayward J, Chansin G, Zervos H. Wearable technology 2017-2027: markets, players, forecasts. IDTechEx. 2017. URL: https://www.idtechex.com/en/research-report/wearable-technology-2017-2027-markets-players-forecasts/536 [accessed 2025-03-17]
9. Guu TW, Muurling M, Khan Z, Kalafatis C, Aarsland D, ffytche D, et al. Wearable devices: underrepresentation in the ageing society. Lancet Digit Health 2023 Jun;5(6):e336-e337. [doi: 10.1016/s2589-7500(23)00069-9]
10. Neubauer N, Spenrath C, Philip S, Daum C, Liu L, Miguel-Cruz A. Identifying adoption and usability factors of locator devices for persons living with dementia. Dementia (London) 2022 Apr 29;21(3):862-881 [FREE Full text] [doi: 10.1177/14713012211065381] [Medline: 34964391]
11. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017 Jan 16;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]

12. Zúñiga-Santamaría T, Vázquez-García CJ. Ethical issues in research in aging. In: García-Peña C, Pérez-Zepeda MU, Gutiérrez-Robledo LM, editors. Aging Research - Methodological Issues Second Edition. Cham, Switzerland: Springer; 2018:209-220.
13. Topfer LA. GPS locator devices for people with dementia. In: CADTH Issues in Emerging Health Technologies. Ottawa, ON: Canadian Agency for Drugs and Technologies in Health; 2016.
14. Bantry-White E. Supporting ethical use of electronic monitoring for people living with dementia: social work's role in assessment, decision-making, and review. J Gerontol Soc Work 2018 Apr 16;61(3):261-279. [doi: 10.1080/01634372.2018.1433738] [Medline: 29381128]
15. Paolillo EW, Lee SY, VandeBunte A, Djukic N, Fonseca C, Kramer JH, et al. Wearable use in an observational study among older adults: adherence, feasibility, and effects of clinicodemographic factors. Front Digit Health 2022 Jun 10;4:884208. [doi: 10.3389/fdgth.2022.884208] [Medline: 35754462]
16. Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. JMIR Mhealth Uhealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
17. Cho S, Ensari I, Weng C, Kahn MG, Natarajan K. Factors affecting the quality of person-generated wearable device data and associated challenges: rapid systematic review. JMIR Mhealth Uhealth 2021 Mar 19;9(3):e20738 [FREE Full text] [doi: 10.2196/20738] [Medline: 33739294]
18. Neubauer NA, Liu L. Influence of perspectives on user adoption of wander-management strategies. Dementia (London) 2021 Feb 12;20(2):734-758. [doi: 10.1177/1471301220911304] [Medline: 32164446]
19. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. JMIR Mhealth Uhealth 2021 Jun 03;9(6):e23832 [FREE Full text] [doi: 10.2196/23832] [Medline: 34081020]
20. Favela J, Cruz-Sandoval D, Morales-Tellez A, Lopez-Nava IH. Monitoring behavioral symptoms of dementia using activity trackers. J Biomed Inform 2020 Sep;109:103520 [FREE Full text] [doi: 10.1016/j.jbi.2020.103520] [Medline: 32783922]
21. Kamil RJ, Bakar D, Ehrenburg M, Wei EX, Pletnikova A, Xiao G, et al. Detection of wandering behaviors using a body-worn inertial sensor in patients with cognitive impairment: a feasibility study. Front Neurol 2021;12:529661 [FREE Full text] [doi: 10.3389/fneur.2021.529661] [Medline: 33776875]
22. Buckley C, Cavadino A, Del Din S, Lord S, Taylor L, Rochester L, et al. Quantifying reliable walking activity with a wearable device in aged residential care: how many days are enough? Sensors (Basel) 2020 Nov 05;20(21):6314 [FREE Full text] [doi: 10.3390/s20216314] [Medline: 33167527]
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
24. Morse JM, Field PA. An overview of qualitative methods. In: Nursing Research. Boston, MA: Springer; 1996.
25. Lazarou I, Stavropoulos TG, Meditskos G, Andreadis S, Kompatsiaris I, Tsolaki M. Long-term impact of intelligent monitoring technology on people with cognitive impairment: an observational study. J Alzheimers Dis 2019 Aug 03;70(3):757-792. [doi: 10.3233/jad-190423]
26. Lee M, Mishra RK, Momin A, El-Refaei N, Bagheri AB, York MK, et al. Smart-home concept for remote monitoring of instrumental activities of daily living (IADL) in older adults with cognitive impairment: a proof of concept and feasibility study. Sensors (Basel) 2022 Sep 07;22(18):6745 [FREE Full text] [doi: 10.3390/s22186745] [Medline: 36146095]
27. Mishra RK, Park C, Momin AS, Rafaei NE, Kunik M, York MK, et al. Care4AD: a technology-driven platform for care coordination and management: acceptability study in dementia. Gerontology 2023 Sep 12;69(2):227-238 [FREE Full text] [doi: 10.1159/000526219] [Medline: 36096091]
28. Amato F, Bianchi S, Comai S, Crovari P, Pasquarelli MG, Imtiaz A, et al. CLONE: a promising system for the remote monitoring of Alzheimer's patients: an experimentation with a wearable device in a village for Alzheimer's care. In: Proceedings of the 4th EAI International Conference on Smart Objects and Technologies for Social Good. 2018 Presented at: Goodtechs '18; November 28-30, 2018; Bologna, Italy. [doi: 10.1145/3284869.3284906]
29. Cruz-Sandoval D, Favela J, Lopez-Nava IH, Morales A. Adoption of wearable devices by persons with dementia: lessons from a non-pharmacological intervention enabled by a social robot. In: Marques G, Bhoi AK, de Albuquerque VH, Hareesha KS, editors. IoT in Healthcare and Ambient Assisted Living. Singapore, Singapore: Springer; 2021:145-163.
30. Engelsma T, Yurt A, Dröes RM, Jaspers MW, Peute LW. Expert appraisal and prioritization of barriers to mHealth use for older adults living with Alzheimer's disease and related dementias: a Delphi study. Int J Med Inform 2022 Oct;166:104845 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104845] [Medline: 35973365]
31. Jacklin K, Pitawanakwat K, Blind M, Lemieux AM, Sobol A, Warry W. Peace of mind: a community-industry-academic partnership to adapt dementia technology for Anishinaabe communities on Manitoulin Island. J Rehabil Assist Technol Eng 2020 Sep 16;7:2055668320958327 [FREE Full text] [doi: 10.1177/2055668320958327] [Medline: 32999732]
32. De Vito AN, Sawyer RJ2, LaRoche A, Arredondo B, Mizuki B, Knoop C. Acceptability and feasibility of a multicomponent telehealth care management program in older adults with advanced dementia in a residential memory care unit. Gerontol Geriatr Med 2020;6:2333721420924988 [FREE Full text] [doi: 10.1177/2333721420924988] [Medline: 32577434]

33. Gelonch O, Ribera M, Codern-Bové N, Ramos S, Quintana M, Chico G, et al. Acceptability of a lifelogging wearable camera in older adults with mild cognitive impairment: a mixed-method study. BMC Geriatr 2019 Apr 16;19(1):110 [FREE Full text] [doi: 10.1186/s12877-019-1132-0] [Medline: 30991948]
34. Godkin FE, Turner E, Demnati Y, Vert A, Roberts A, Swartz RH, et al. Feasibility of a continuous, multi-sensor remote health monitoring approach in persons living with neurodegenerative disease. J Neurol 2022 May 27;269(5):2673-2686 [FREE Full text] [doi: 10.1007/s00415-021-10831-z] [Medline: 34705114]
35. Hsu CK, Liu CC, Chang T, Liao JJ, Shu CM. Service design of a loss prevention device for older adults with dementia. Geriatrics (Basel) 2023 Sep 15;8(5):93 [FREE Full text] [doi: 10.3390/geriatrics8050093] [Medline: 37736893]
36. Anderson MS, Bankole A, Homdee N, Mitchell BA, Byfield GE, Lach J. Dementia caregiver experiences and recommendations for using the behavioral and environmental sensing and intervention system at home: usability and acceptability study. JMIR Aging 2021 Dec 06;4(4):e30353 [FREE Full text] [doi: 10.2196/30353] [Medline: 34874886]
37. Dai B, Larnyo E, Tetteh EA, Aboagye AK, Musah AI. Factors affecting caregivers' acceptance of the use of wearable devices by patients with dementia: an extension of the unified theory of acceptance and use of technology model. Am J Alzheimers Dis Other Demen 2020 Nov 03;35:1533317519883493 [FREE Full text] [doi: 10.1177/1533317519883493] [Medline: 31679390]
38. Gris F, D'Amen B, Lamura G, Paciaroni L, Socci M, Melchiorre MG. Personalized technological support for informal caregivers of older people with dementia: a co-design approach involving potential end users and healthcare professionals in three focus groups in Italy. Healthcare (Basel) 2023 Sep 28;11(19):2640 [FREE Full text] [doi: 10.3390/healthcare11192640] [Medline: 37830677]
39. Larnyo E, Dai B, Larnyo A, Nutakor JA, Ampon-Wireko S, Nkrumah EN, et al. Impact of actual use behavior of healthcare wearable devices on quality of life: a cross-sectional survey of people with dementia and their caregivers in Ghana. Healthcare (Basel) 2022 Jan 30;10(2):275 [FREE Full text] [doi: 10.3390/healthcare10020275] [Medline: 35206890]
40. Ahmed QA, Al-Neami AQ. A smart biomedical assisted system for Alzheimer patients. IOP Conf Ser Mater Sci Eng 2020 Jul 01;881:012110 [FREE Full text] [doi: 10.1088/1757-899X/881/1/012110]
41. Farina N, Sherlock G, Thomas S, Lowry RG, Banerjee S. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Int J Geriatr Psychiatry 2019 Apr 20;34(4):617-624. [doi: 10.1002/gps.5064] [Medline: 30701592]
42. Megges H, Freiesleben SD, Rösch C, Knoll N, Wessel L, Peters O. User experience and clinical effectiveness with two wearable global positioning system devices in home dementia care. Alzheimers Dement (N Y) 2018;4:636-644 [FREE Full text] [doi: 10.1016/j.trci.2018.10.002] [Medline: 30519629]
43. McCarron HR, Zmora R, Gaugler JE. A web-based mobile app with a smartwatch to support social engagement in persons with memory loss: pilot randomized controlled trial. JMIR Aging 2019 Jun 18;2(1):e13378 [FREE Full text] [doi: 10.2196/13378] [Medline: 31518270]
44. Freytag J, Mishra RK, Street RLJ, Catic A, Dindo L, Kiefer L, et al. Using wearable sensors to measure goal achievement in older veterans with dementia. Sensors (Basel) 2022 Dec 16;22(24):9923 [FREE Full text] [doi: 10.3390/s22249923] [Medline: 36560290]
45. Mc Ardle R, Del Din S, Donaghy P, Galna B, Thomas AJ, Rochester L. The impact of environment on gait assessment: considerations from real-world gait analysis in dementia subtypes. Sensors (Basel) 2021 Jan 26;21(3):813 [FREE Full text] [doi: 10.3390/s21030813] [Medline: 33530508]
46. Hall A, Brown Wilson C, Stanmore E, Todd C. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. BMC Geriatr 2019 May 24;19(1):145 [FREE Full text] [doi: 10.1186/s12877-019-1155-6] [Medline: 31126240]
47. Guu TW, Aarsland D, Ffytche DH. Feasibility of using wearable actigraphy for behavioural monitoring in care home residents living with moderate to severe Alzheimer's disease and agitation. Alzheimers Dement 2023 Dec 25;19(S19):e072283. [doi: 10.1002/alz.072283]
48. Musaeus CS, Waldemar G, Andersen BB, Høgh P, Kidmose P, Hemmsen MC, et al. Long-term EEG monitoring in patients with Alzheimer's disease using Ear-EEG: a feasibility study. J Alzheimers Dis 2022 Dec 06;90(4):1713-1723. [doi: 10.3233/jad-220491]
49. Dinesh K, Snyder CW, Xiong M, Tarolli CG, Sharma S, Dorsey ER, et al. A longitudinal wearable sensor study in Huntington's disease. J Huntingtons Dis 2020;9(1):69-81. [doi: 10.3233/JHD-190375] [Medline: 31868675]
50. Mc Ardle R, Morris R, Hickey A, Del Din S, Koychev I, Gunn RN, et al. Gait in mild Alzheimer's disease: feasibility of multi-center measurement in the clinic and home with body-worn sensors: a pilot study. J Alzheimers Dis 2018 Apr 10;63(1):331-341. [doi: 10.3233/jad-171116]
51. Berridge C, Zhou Y, Lazar A, Porwal A, Mattek N, Gothard S, et al. Control matters in elder care technology: evidence and direction for designing it in. DIS (Des Interact Syst Conf) 2022 Jun;2022:1831-1848 [FREE Full text] [doi: 10.1145/3532106.3533471] [Medline: 35969716]
52. Nickerson GY, Shade K. Effect of increasing physical activity on cognitive function in individuals with mild cognitive impairment: a knowledge translation to practice pilot project. J Gerontol Nurs 2021 May;47(5):14-18 [FREE Full text] [doi: 10.3928/00989134-20210401-01] [Medline: 34039095]

53. Cohen S, Waks Z, Elm JJ, Gordon MF, Grachev ID, Navon-Perry L, et al. Characterizing patient compliance over six months in remote digital trials of Parkinson's and Huntington disease. BMC Med Inform Decis Mak 2018 Dec 20;18:138. [doi: 10.1186/s12911-018-0714-7]
54. Eckert T, Bongartz M, Ullrich P, Abel B, Christian W, Kiss R, et al. Promoting physical activity in geriatric patients with cognitive impairment after discharge from ward-rehabilitation: a feasibility study. Eur J Ageing 2020 Sep 13;17(3):309-320 [FREE Full text] [doi: 10.1007/s10433-020-00555-w] [Medline: 32904844]
55. Chen JH, Lauderdale DS. Cognitive function, consent for participation, and compliance with wearable device protocols in older adults. J Gerontol A Biol Sci Med Sci 2019 Jan 16;74(2):269-273 [FREE Full text] [doi: 10.1093/gerona/gly032] [Medline: 29579176]
56. Dillon K, Prapavessis H. REducing SEDENTary behavior among mild to moderate cognitively impaired assisted living residents: a pilot randomized controlled trial (RESEDENT study). J Aging Phys Act 2021 Feb 01;29(1):27-35 [FREE Full text] [doi: 10.1123/japa.2019-0440] [Medline: 32580164]
57. Gibson RH, Gander PH. Monitoring the sleep patterns of people with dementia and their family carers in the community. Australas J Ageing 2019 Mar 07;38(1):47-51. [doi: 10.1111/ajag.12605] [Medline: 30525266]
58. Bowen ME, Gaynor B, Phillips LJ. Changes in physical and cognitive function predict sedentary behavior in older adults with mild cognitive impairment. Res Gerontol Nurs 2021 Nov;14(6):285-291. [doi: 10.3928/19404921-20211021-01] [Medline: 34807787]
59. Gedde MH, Husebo BS, Erdal A, Puaschitz NG, Vislapuu M, Angeles RC, et al. Access to and interest in assistive technology for home-dwelling people with dementia during the COVID-19 pandemic (PAN.DEM). Int Rev Psychiatry 2021 Jun 08;33(4):404-411 [FREE Full text] [doi: 10.1080/09540261.2020.1845620] [Medline: 33416012]
60. Kaenampornpan M, Khai ND, Kawattikul K. Wearable computing for dementia patients. In: Proceedings of the 16th International Conference on Computing and Information Technology. 2020 Presented at: IC2IT 2020; May 14-15, 2020; Pattaya, Thailand. [doi: 10.1007/978-3-030-44044-2_3]
61. Peeters MW, Schouten G, Wouters EJ. Wearables for residents of nursing homes with dementia and challenging behaviour: values, attitudes, and needs. Gerontechnology 2021 Jan 01;20(2):1-13. [doi: 10.4017/gt.2021.20.2.7.06]
62. Richeson NE, Croteau KA. A feasibility study examining use of the FitBit Zip™ vs. the Accusplit Eagle AC 120 XL pedometer to increase physical activity for persons with mild cognitive disorder. Activ Adapt Aging 2017 Oct 16;42(1):41-53. [doi: 10.1080/01924788.2017.1385367]
63. Stavropoulos TG, Lazarou I, Diaz A, Gove D, Georges J, Manyakov NV, et al. Wearable devices for assessing function in Alzheimer's disease: a European public involvement activity about the features and preferences of patients and caregivers. Front Aging Neurosci 2021 Apr 12;13:643135 [FREE Full text] [doi: 10.3389/fnagi.2021.643135] [Medline: 33912025]
64. Stavropoulos TG, Lazarou I, Strantsalis D, Nikolopoulos S, Kompatsiaris I, Koumanakos G. Human factors and requirements of people with mild cognitive impairment, their caregivers and healthcare professionals for eHealth systems with wearable trackers. In: Proceedings of the 2020 IEEE International Conference on Human-Machine Systems. 2020 Presented at: ICHMS 2020; September 7-9, 2020; Rome, Italy. [doi: 10.1109/ichms49158.2020.9209340]
65. van der Wardt V, Hancox JE, Burgon C, Bajwa R, Goldberg S, Harwood RH. Measuring physical activity levels in people with mild cognitive impairment or mild dementia. J Aging Phys Act 2021 Feb 01;29(1):10-16. [doi: 10.1123/japa.2019-0234] [Medline: 33049697]
66. Rose KM, Lach J, Perkhounkova Y, Gong J, Dandu SR, Dickerson R, et al. Use of body sensors to examine nocturnal agitation, sleep, and urinary incontinence in individuals with Alzheimer's disease. J Gerontol Nurs 2018 Aug 01;44(8):19-26. [doi: 10.3928/00989134-20180626-03] [Medline: 30059136]
67. Thorpe J, Forchhammer BH, Maier AM. Adapting mobile and wearable technology to provide support and monitoring in rehabilitation for dementia: feasibility case series. JMIR Form Res 2019 Oct 17;3(4):e12346 [FREE Full text] [doi: 10.2196/12346] [Medline: 31625951]
68. Wherton J, Greenhalgh T, Procter R, Shaw S, Shaw J. Wandering as a sociomaterial practice: extending the theorization of GPS tracking in cognitive impairment. Qual Health Res 2019 Feb 14;29(3):328-344 [FREE Full text] [doi: 10.1177/1049732318798358] [Medline: 30215572]
69. Raepsaet C, Serraes B, Verhaeghe S, Beeckman D. Integrating sensor technology in disposable body-worn absorbent products: a qualitative study to define user profile, (technical) criteria, conditions, and potential benefits. J Wound Ostomy Continence Nurs 2021;48(6):560-567. [doi: 10.1097/WON.0000000000000812] [Medline: 34781313]
70. Parry S, Chow M, Batchelor F, Fary RE. Physical activity and sedentary behaviour in a residential aged care facility. Australas J Ageing 2019 Mar 03;38(1):E12-E18. [doi: 10.1111/ajag.12589] [Medline: 30281184]
71. Snyder M, Dringus L, Maitland Schladen M, Chenail R, Oviawe E. Remote monitoring technologies in dementia care: an interpretative phenomenological analysis of family caregivers' experiences. Qual Report 2020 May 11;25(5):1233-1252. [doi: 10.46743/2160-3715/2020.4127]
72. Sharma N, Braakman-Jansen LM, Oinas-Kukkonen H, Croockewit JH, van Gemert-Pijnen J. Exploring the needs and requirements of informal caregivers of older adults with cognitive impairment from sensor-based care solutions: multimethod study. JMIR Aging 2023 Oct 25;6:e49319 [FREE Full text] [doi: 10.2196/49319] [Medline: 37878353]

73. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med Ethics 2019 Dec 19;20(1):98 [FREE Full text] [doi: 10.1186/s12910-019-0437-z] [Medline: 31856798]
74. Tiersen F, Batey P, Harrison MJ, Naar L, Serban AI, Daniels SJ, et al. Smart home sensing and monitoring in households with dementia: user-centered design approach. JMIR Aging 2021 Aug 11;4(3):e27047 [FREE Full text] [doi: 10.2196/27047] [Medline: 34383672]
75. Svetnik V, Wang TC, Ceesay P, Snyder E, Ceren O, Bliwise D, et al. Pilot evaluation of a consumer wearable device to assess sleep in a clinical polysomnography trial of suvorexant for treating insomnia in patients with Alzheimer's disease. J Sleep Res 2021 Dec 02;30(6):e13328. [doi: 10.1111/jsr.13328] [Medline: 34340251]
76. Sun Y, Kim HM, Xu Y, Wang Y, Kwong J, Kim SS, et al. GPS tracking in dementia caregiving: social norm, perceived usefulness, and behavioral intent to use technology. In: Proceedings of the 54th Hawaii International Conference on System Sciences. 2021 Presented at: HICSS 2021; January 5, 2021; Kauai, HI. [doi: 10.24251/hicss.2021.461]
77. O'Sullivan G, Whelan B, Gallagher N, Doyle P, Smyth S, Murphy K, et al. Challenges of using a Fitbit smart wearable among people with dementia. Int J Geriatr Psychiatry 2023 Mar 28;38(3):e5898. [doi: 10.1002/gps.5898] [Medline: 36814072]
78. Hansson SO. Technology and distributive justice. In: The Ethics of Technology: Methods and Approaches. Lanham, MD: Rowman & Littlefield Publishers; 2017:51-66.
79. Guu TW, Brem AK, Albertyn CP, Kandangwa P, Aarsland D, Ffytche D. Wrist-worn actigraphy in agitated late-stage dementia patients: a feasibility study on digital inclusion. Alzheimers Dement 2024 May 18;20(5):3211-3218 [FREE Full text] [doi: 10.1002/alz.13772] [Medline: 38497216]
80. Goerss D, Köhler S, Rong E, Temp AG, Kilimann I, Bieber G, et al. Smartwatch-based interventions for people with dementia: user-centered design approach. JMIR Aging 2024 Jun 07;7:e50107 [FREE Full text] [doi: 10.2196/50107] [Medline: 38848116]
81. Muurling M, de Boer C, Hinds C, Atreya A, Doherty A, Alepopoulos V, et al. Feasibility and usability of remote monitoring in Alzheimer's disease. Digit Health 2024 Apr 09;10:20552076241238133 [FREE Full text] [doi: 10.1177/20552076241238133] [Medline: 38601188]
82. Gaugler JE, Zmora R, Mitchell LL, Finlay J, Rosebush CE, Nkimbeng M, et al. Remote activity monitoring for family caregivers of persons living with dementia: a mixed methods, randomized controlled evaluation. BMC Geriatr 2021 Dec 18;21(1):715 [FREE Full text] [doi: 10.1186/s12877-021-02634-8] [Medline: 34922475]
83. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
84. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Ind Manag Data Syst 2015 Oct 19;115(9):1704-1723. [doi: 10.1108/IMDS-03-2015-0087]
85. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008 May 09;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
86. Use of wearable fitness or wellness technologies in the United States as of 2023, by age. Statista. URL: https://www.statista.com/statistics/1424304/use-of-fitness-wearables-in-the-us-by-age/ [accessed 2024-06-21]
87. Chandrasekaran R, Katthula V, Moustakas E. Patterns of use and key predictors for the use of wearable health care devices by US adults: insights from a national survey. J Med Internet Res 2020 Oct 16;22(10):e22443 [FREE Full text] [doi: 10.2196/22443] [Medline: 33064083]
88. Chinner A, Blane J, Lancaster C, Hinds C, Koychev I. Digital technologies for the assessment of cognition: a clinical review. Evid Based Ment Health 2018 May 20;21(2):67-71 [FREE Full text] [doi: 10.1136/eb-2018-102890] [Medline: 29678927]
89. Winer JR, Lok R, Weed L, He Z, Poston KL, Mormino EC, et al. Impaired 24-h activity patterns are associated with an increased risk of Alzheimer's disease, Parkinson's disease, and cognitive decline. Alzheimers Res Ther 2024 Feb 14;16(1):35 [FREE Full text] [doi: 10.1186/s13195-024-01411-0] [Medline: 38355598]

Abbreviations

MeSH: Medical Subject Headings
PICO: population, intervention, comparison, and outcomes
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUEST: Quebec User Evaluation of Satisfaction With Assistive Technology

Edited by M Mardini; submitted 28.06.24; peer-reviewed by TW Guu, M Gagnon-Roy; comments to author 04.02.25; revised version received 25.03.25; accepted 27.04.25; published 31.07.25.

Please cite as:
Peterson CM, St Louis RM, Flannagan C
Enhancing Enrollment and Adherence in Long-Term Wearable Research on Dementia: Qualitative Systematic Review and Meta-Synthesis
JMIR Aging 2025;8:e63768
URL: https://aging.jmir.org/2025/1/e63768
doi:10.2196/63768
PMID:

Viewpoint

A Digitally Capable Aged Care Workforce: Demands and Directions for Workforce Education and Development

Kathleen Gray[1], PhD; Kerryn Butler-Henderson[2], PhD; Karen Day[3], PhD

[1]Centre for Digital Transformation of Health, University of Melbourne, Melbourne, Australia

[2]School of Nursing, Paramedicine, and Healthcare Sciences, Charles Sturt University, Bathurst, Australia

[3]Health Systems Group, University of Auckland, Auckland, New Zealand

Corresponding Author:
Kathleen Gray, PhD
Centre for Digital Transformation of Health
University of Melbourne
Parkville
Melbourne, 3010
Australia
Phone: 61 390355511
Email: kgray@unimelb.edu.au

Abstract

As the aged care sector undergoes digital transformation, greater attention is needed to development of digital health capability in its workforce. There are many gaps in our understanding of the current and future impacts of technology on those who perform paid and unpaid aged care work. Research is needed to understand how to make optimal use of both digital resources and human resources for better aged care. In this Viewpoint, we reflect on a workshop held during an international conference that identified shared concepts and concerns to shape further research into workforce capability. Digital technologies and digital data can increase quality of care in a system that operates through partnerships among service providers, service users, and community members. To realize this potential, digital health learning and development are needed in the aged care workforce. As digital dimensions of aged care services expand, the sector needs clearer direction to implement approaches to workforce learning and development. These must be appropriate to support the safe and ethical performance of care work and to increase the satisfaction of those who care and those for whom they care.

(JMIR Aging 2025;8:e54143) doi:10.2196/54143

KEYWORDS

aged care; digital health; digital literacy; education; older adults; professional development; digital transformation; digital resources; users; community; learning; support; safe; ethical; satisfaction

Why Focus on a Digitally Capable Aged Care Workforce?

The aged care sector is in the spotlight around the world due to the increasing health and well-being needs of aging populations. Examples are the Australian Royal Commission into Aged Care Quality and Safety [1], Japan's Social Welfare for the Elderly Act [2], and Germany's national report examining the human rights of older persons in long-term care [3]. Comparisons of aged care systems internationally focus on a variety of issues [4,5].

In such reports, the aged care workforce emerges as an essential factor. One part of this workforce is paid, structured, and regulated (eg, registered nurses) and formally counted in the health and care workforce. Another part of this workforce is informal—often unpaid and not structured or regulated (eg, relatives as unpaid caregivers) but contributing in incalculable ways [6]. Both parts operate under conditions that can be inequitable, inefficient, uncertain, and unsatisfactory. The overall situation of the aged care workforce demands widespread improvements in recruitment and retention, practice standards, skill levels, recognition, and reward [7].

Digital maturity in aged care service provider organizations lags behind mainstream health care services; there are differences between the two also in terms of information needs, technologies of interest, and administration. The medico-legal record for aged care services works differently from hospitals and primary care services, making data and information standardization problematic [8]. Robotics, sensors, and virtual reality have been

explored for over a decade in many countries for their potential to augment human care-giving, but with little opportunity for uptake due to cost and complexity [9]. Telehealth has been shown to have a promising place in aged care, especially since the COVID-19 pandemic, but funding and service supply remain problematic [10]. Nevertheless, digital transformation of the aged care sector is underway [11,12].

Despite this, digital skills of any kind rarely appear in aged care job descriptions or training programs. Further, national workforce strategies related to digital health capability scarcely address the digital skill shortages in the aged care sector. On one hand, there is hope that the sector will find the "right" staff for digitally enabled aged care and find ways to assure them of technology's quality and trustworthiness [13,14]. On the other hand, there is uneasiness that roles and responsibilities may become invisible, peripheral, or "care-adjacent" in the process of moving to technological efficiency and sustainability [15,16]. In short, the introduction of technology has still-unknown implications for the evolution of work in the aged care sector, with possible effects ranging from displacing workers to supporting productivity in current roles or to elevating responsibilities and recognition. However, much of the research to date into aged carers' digital experiences and expectations has had a short-term focus on technology implementations rather than consequences and outcomes [17,18].

Working with technology in particular ways can support specific well-being factors, including positive emotions, self-awareness, mindfulness, empathy, and compassion—so-called positive computing [19]—and there are reports that this approach can benefit aged care organizations and their workers [20]. But there is no guarantee that this will happen without a deliberate workforce strategy. This raises questions about how to proceed, which shape the aims of further research: How can digital transformation of the aged care sector be optimized through greater investment in the digital knowledge, skills, attributes, and attitudes of its workforce? How can such investment in turn improve working conditions and job satisfaction in the aged care workforce?

The authors explored these questions during a workshop that we facilitated at the July 2023 MedInfo conference in Sydney, Australia. There we presented our view that workforce research and development is needed to achieve optimal outcomes for workers in order to achieve these for service users as the aged care sector moves toward digital maturity. The ensuing discussions with workshop delegates contributed to refining the viewpoint we outline here.

What Knowledge and Skill Building Is Required for Digital Aged Care Work?

Digital aged care work is likely to deliver the greatest benefits when needs for fundamental, complementary digital knowledge and skills are met in both paid and unpaid workers. Table 1 summarizes and juxtaposes the capabilities that we propose.

Table 1. Digital capabilities needed in the aged care workforce.

For paid work	For unpaid work
Knowing about developments in the digital environment in which health and care operate	Knowing about the existence of digital systems and tools, and the ways they can support care users and workers
Knowing how to adopt digital tools to help with care coordination and monitoring	Knowing how to use tools to support care closer to home and enhance care provision
Knowing the importance for care quality of collecting and interpreting digital data collected during episodes of care	Knowing how to navigate a complex aged care system across multiple digital platforms
Ability with general digital proficiency and literacy, including information security	Ability with general digital proficiency and literacy
Ability to evaluate online information critically	Ability to access and interpret information from multiple online sources
Ability to consider usability and empathy when using digital tools with patients and clients	Ability to learn to engage with technology and be comfortable with technology
Ability to interpret biometric data and help the patient or client to understand it	Ability to be connected and share health data between care consumers and providers
Ability to judge when technology is not needed as well as when it is	Ability to access and use technology to overcome isolation

There are parallel aspects of capability, with different levels of sophistication, in the needs of paid and unpaid workers. Accordingly, paid workers must be cognizant of not only their own capabilities, but also the capabilities of the unpaid workers with whom they partner. Aside from the digital capabilities of direct consumers of aged care services, paid workers must consider and collaborate with the capabilities of volunteer carers among family, friends and community members. Thus, to build the digitally capable human resources that will support healthy and dignified aging for all, the sector certainly must increase levels of digital literacy in the paid workforce; digital literacy being defined as the ability to use technology to communicate, retrieve, and evaluate information to make decisions [21]. However, the sector also must increase levels of active engagement in digital citizenship among both paid and unpaid workers; in the concept of digital citizenship, people are increasingly expected to use a range of technologies to access services of all kinds and to participate on all levels in society [22].

Various approaches are possible to support digital knowledge building and skill building for the paid workforce. There seem to be gaps and requirements in terms of both formally accredited

training programs and less formal continuing professional development, as well as for both workplace-based learning and workplace-independent learning. The diverse and differentiated methods needed to uplift the digital capabilities of the unpaid workforce are an even broader challenge. Strategy and culture will have to change in the aged care sector to create an environment that encourages and supports all those who do the work to embrace learning and development for positive uses of technology. Among service providers, technology suppliers, industry and community associations, and government agencies, it is not yet clear where roles and responsibilities should sit for planning, implementing, and evaluating the changes that are needed to create and support a digitally capable aged care workforce.

Although these themes are similar to those arising in the mainstream health care workforce in some respects, they also recognize that the aged care ecosystem is distinctive and heterogeneous in terms of the array of partnerships among service providers, service users, and community members [23]. It may be useful to build a more detailed understanding of the digital capability gaps and needs in this sector by overlaying two other system-level concepts onto the idea of the aged care ecosystem—care ecosystems and workforce ecosystems. A care ecosystem is defined as an approach to configuring paid and unpaid workers, ambient assistive technologies, and digital data sharing as work-arounds for intractable workforce shortages in care, particularly in situations that do not call for advanced or intensive clinical services [24,25]. A workforce ecosystem is characterized by employees and contractors working interdependently with automated processes to achieve the goals of an organization and their personal goals; it is considered a guiding concept for the challenges of human resources management as artificial intelligence changes industries and societies [26,27]. A whole-system approach that acknowledges these concepts could provide a coherent framework to plan, implement, and evaluate the range of complementary measures that are needed for digitally enabled aged care work.

Conclusions and Next Steps

Digital technologies and digital data have the potential to improve many aspects of the experiences of the people who receive aged care and the people who provide it. The context in which we seek to develop the digital capabilities of the aged care workforce is challenging. There are strong sustainability pressures on service providers. Public policies, strategies, and sentiments favor innovative, distributed care. Social and technological services and solutions are not joined up. Workers across the sector are stretched thin, whether they are highly trained professionals (eg, registered nurses), minimally trained caregivers (eg, health and personal care assistants), or informal caregivers (eg families, friends, and neighbours).

Understanding how to build digital capability in this workforce is complex and multifaceted, with many unknowns. Inclusion is one: What learning and development methods are most efficient and effective to enable digital skills across the entirety of this workforce? Empowerment is another: How can aged care workers be enabled to advocate for and implement digital improvements that bring positive changes to care? Expertise is another: What career opportunities and leadership roles are appropriate to motivate digital aged care specialists and sustain their attention to workforce development?

Next steps for the work in this field include reviewing the literature to distill evidence from reports of small-scale efforts, workforce research, and development that engages real-world settings and stakeholders, as well as trials of scalable learning and professional and community development programs. To progress this research requires new collaborations among aged care workers, service provider organizations, aged care technology vendors, and digital health workforce researchers. We can improve the current standing of digital skills in the aged care workforce only if we first clarify what the workforce needs from education providers, employers, technology suppliers, and policy makers to enable and motivate them to use existing and emerging technologies safely and wisely. This work will provide important information to design learning and development that enhances digital capability, and it will establish expectations that allow us to evaluate changes in workforce status and performance as an integral part of the digital transformation of aged care.

Acknowledgments

We are grateful to the convenors of the MedInfo 2023 conference for the opportunity to conduct our workshop and to the delegates who attended and engaged in discussion. The 3 authors prepared the workshop together. To set up the workshop discussion, each author presented on a viewpoint of the topic, followed by KB-H and KD cochairing the discussion. All 3 authors contributed to writing and editing this Viewpoint and signed off the paper as ready for submission.

Conflicts of Interest

None declared.

References

1. Final report: care, dignity and response. Royal Commission into Aged Care Quality and Safety. 2021. URL: https://www.royalcommission.gov.au/aged-care/final-report [accessed 2025-03-28]

2. Nakamura S. Japan's welfare system for the elderly: past, present, and future. Asia Health and Wellbeing Initiative. URL: https://www.ahwin.org/wp-content/uploads/2020/06/201812-S-Nakamura-Japan%E2%80%99s-Welfare-System-for-the-Elderly-.pdf [accessed 2025-03-28]
3. Aronson P, Mahler C. Human rights of older persons in long-term care: German national report. European Network of National Human Rights Institutions. 2016. URL: https://ennhri.org/wp-content/uploads/2019/10/germany.pdf [accessed 2025-03-28]
4. Review of international systems for long-term care of older people. Research paper 2. Royal Commission into Aged Care Quality and Safety. 2020. URL: https://apo.org.au/node/274151 [accessed 2025-04-02]
5. Appendix D: international experience – inquiry report – caring for older Australians. Australian Government Productivity Division. 2011. URL: https://www.pc.gov.au/inquiries/completed/aged-care/report/26-aged-care-appendixd.pdf [accessed 2025-03-28]
6. Woolford M, Joyce C, Polacsek M. Strategies to optimize aged care volunteer recruitment and retention: a systematic review. J Gerontol Soc Work 2022 Jan;65(1):45-62. [doi: 10.1080/01634372.2021.1923605] [Medline: 33971797]
7. Who cares? Attracting and retaining care workers for the elderly. Organisation for Economic Co-operation and Development. 2020. URL: https://www.oecd.org/en/publications/who-cares-attracting-and-retaining-elderly-care-workers_92c0ef68-en.html [accessed 2025-03-28]
8. Bail K, Gibson D, Acharya P, Blackburn J, Kaak V, Kozlovskaia M, et al. Using health information technology in residential aged care homes: an integrative review to identify service and quality outcomes. Int J Med Inform 2022 Sep;165:104824 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104824] [Medline: 35792376]
9. Penteridis L, D'Onofrio G, Sancarlo D, Giuliani F, Ricciardi F, Cavallo F, et al. Robotic and sensor technologies for mobility in older people. Rejuvenation Res 2017 Oct;20(5):401-410. [doi: 10.1089/rej.2017.1965] [Medline: 28482748]
10. Day K, Millner S, Johnson H. How nurses use telehealth to support health transitions of older adults. Stud Health Technol Inform 2016;231:23-30. [doi: 10.3233/978-1-61499-712-2-23] [Medline: 27782013]
11. Nordin S, Sturge J, Ayoub M, Jones A, McKee K, Dahlberg L, et al. The role of information and communication technology (ICT) for older adults' decision-making related to health, and health and social care services in daily life-a scoping review. Int J Environ Res Public Health 2021 Dec 23;19(1):151 [FREE Full text] [doi: 10.3390/ijerph19010151] [Medline: 35010408]
12. Fasoli A, Beretta G, Pravettoni G, Sanchini V. Mapping emerging technologies in aged care: results from an in-depth online research. BMC Health Serv Res 2023 May 23;23(1):528 [FREE Full text] [doi: 10.1186/s12913-023-09513-5] [Medline: 37221528]
13. Peters M, Marnie C. Human costs of aged care productivity: innovation versus staffing and skills mix. Collegian 2021 Jun;28(3):351-355 [FREE Full text] [doi: 10.1016/j.colegn.2020.11.002]
14. Wang Y, Feng Y, Liang C. How digital change and innovation in the workplace affect front-line employee retention: a cross-sectional study based on the aged care industry. In: Tu Y, Chi M, editors. E-Business. Digital Empowerment for an Intelligent Future. WHICEB 2023. Lecture Notes in Business Information Processing, vol 480. Cham, Switzerland: Springer; 2023:72-83.
15. Hamblin KA. Technology in care systems: displacing, reshaping, reinstating or degrading roles? New Technol Work Employ 2022 Mar;37(1):41-58 [FREE Full text] [doi: 10.1111/ntwe.12229] [Medline: 35911255]
16. Boeva Y, Berger A, Bischof A, Doggett O, Heuer H, Jarke J, et al. Behind the scenes of automation: ghostly care-work, maintenance, and interferences: exploring participatory practices and methods to uncover the ghostly presence of humans and human labor in automation. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems. New York, NY: Association for Computing Machinery; 2023:1-5.
17. Rigby K, Redley B, Hutchinson AM. Change agent's role in facilitating use of technology in residential aged care: a systematic review. Int J Med Inform 2023 Nov;179:105216 [FREE Full text] [doi: 10.1016/j.ijmedinf.2023.105216] [Medline: 37734272]
18. Bail K, Merrick E, Gibson D, Hind A, Strickland K, Redley B. A co-designed health information system implementation into residential aged care: A mixed-method evaluation. J Nurs Scholarsh 2023 Mar;55(2):521-535. [doi: 10.1111/jnu.12840] [Medline: 36366792]
19. Calvo R, Peters D. Positive Computing: Technology for Wellbeing and Human Potential. Cambridge, MA: MIT Press; 2014.
20. Shearman P. The manifestation of positive computing in the aged care workforce. University of Technology Sydney Faculty of Business [PhD thesis]. 2021. URL: https://opus.lib.uts.edu.au/handle/10453/158673 [accessed 2025-03-28]
21. Nichols T, Stornaiuolo A. Assembling "digital literacies": contingent pasts, possible futures. Media Commun 2019;7(2):14 [FREE Full text] [doi: 10.17645/mac.v7i2.1946]
22. Holgersson J, Söderström E, Rose J. Digital inclusion for elderly citizens for a sustainable society. In: Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden, June 8-14, 2019. Atlanta, GA: Association for Information Systems; 2019.
23. Digital transformation sector partners co-design meeting (2). Department of Health Australia. URL: https://www.health.gov.au/sites/default/files/documents/2022/07/digital-transformation-sector-partners-slide-presentation-6-june-2022.pdf [accessed 2025-03-28]

24. Baldissera TA, Camarinha-Matos LM, De Faveri C. Human factor in designing an elderly care ecosystem. In: Adewumi A, Misra S, editors. Handbook of Research on the Role of Human Factors in IT Project Management. Hershey, PA: IGI Global; 2020:106-131.
25. Tseng H, Huang H, Hsieh C. Active aging AI community care ecosystem design. In: Gao Q, Zhou J, editors. Human Aspects of IT for the Aged Population. Healthy and Active Aging. HCII 2020. Lecture Notes in Computer Science(), vol 12208. Cham, Switzerland: Springer; 2020.
26. Altman E, Kiron D, Schwartz J, Jones R. The future of work is through workforce ecosystems. MIT Sloan Management Review. 2021. URL: https://sloanreview.mit.edu/article/the-future-of-work-is-through-workforce-ecosystems/ [accessed 2025-03-28]
27. Gorski AT, Gligorea I, Gorski H, Oancea R. Workforce and workplace ecosystem – challenges and opportunities in the age of digital transformation and 4IR. Int Conf Knowl Based Org 2022 Jun;28(1):187-194 [FREE Full text] [doi: 10.2478/kbo-2022-0028]

Edited by J Wang; submitted 30.10.23; peer-reviewed by M Kleib, A Azizi; comments to author 23.01.24; revised version received 29.04.24; accepted 12.12.24; published 02.04.25.

Please cite as:
Gray K, Butler-Henderson K, Day K
A Digitally Capable Aged Care Workforce: Demands and Directions for Workforce Education and Development
JMIR Aging 2025;8:e54143
URL: https://aging.jmir.org/2025/1/e54143
doi:10.2196/54143
PMID:

Viewpoint

Advancing Emergency Care With Digital Twins

Haoran Li[1], BSc; Jingya Zhang[1], BSc; Ning Zhang[2], PhD; Bin Zhu[3], PhD

[1]School of Public Policy and Administration, Xi'an Jiaotong University, Xi'an, China

[2]Vanke School of Public Health, Tsinghua University, Beijing, China

[3]School of Public Health and Emergency Management, Southern University of Science and Technology, Shenzhen, China

Corresponding Author:
Bin Zhu, PhD
School of Public Health and Emergency Management
Southern University of Science and Technology
1088 Xueyuan Avenue
Nanshan District
Shenzhen, 518000
China
Phone: 86 13530405020
Email: zhub6@sustech.edu.cn

Abstract

Digital twins—dynamic and real-time simulations of systems or environments—represent a paradigm shift in emergency medicine. We explore their applications across prehospital care, in-hospital management, and recovery. By integrating real-time data, wearable technology, and predictive analytics, digital twins hold the promise of optimizing resource allocation, advancing precision medicine, and tailoring rehabilitation strategies. Moreover, we discuss the challenges associated with their implementation, including data resolution, biological heterogeneity, and ethical considerations, emphasizing the need for actionable frameworks that balance innovation with data governance and public trust.

(JMIR Aging 2025;8:e71777) doi:10.2196/71777

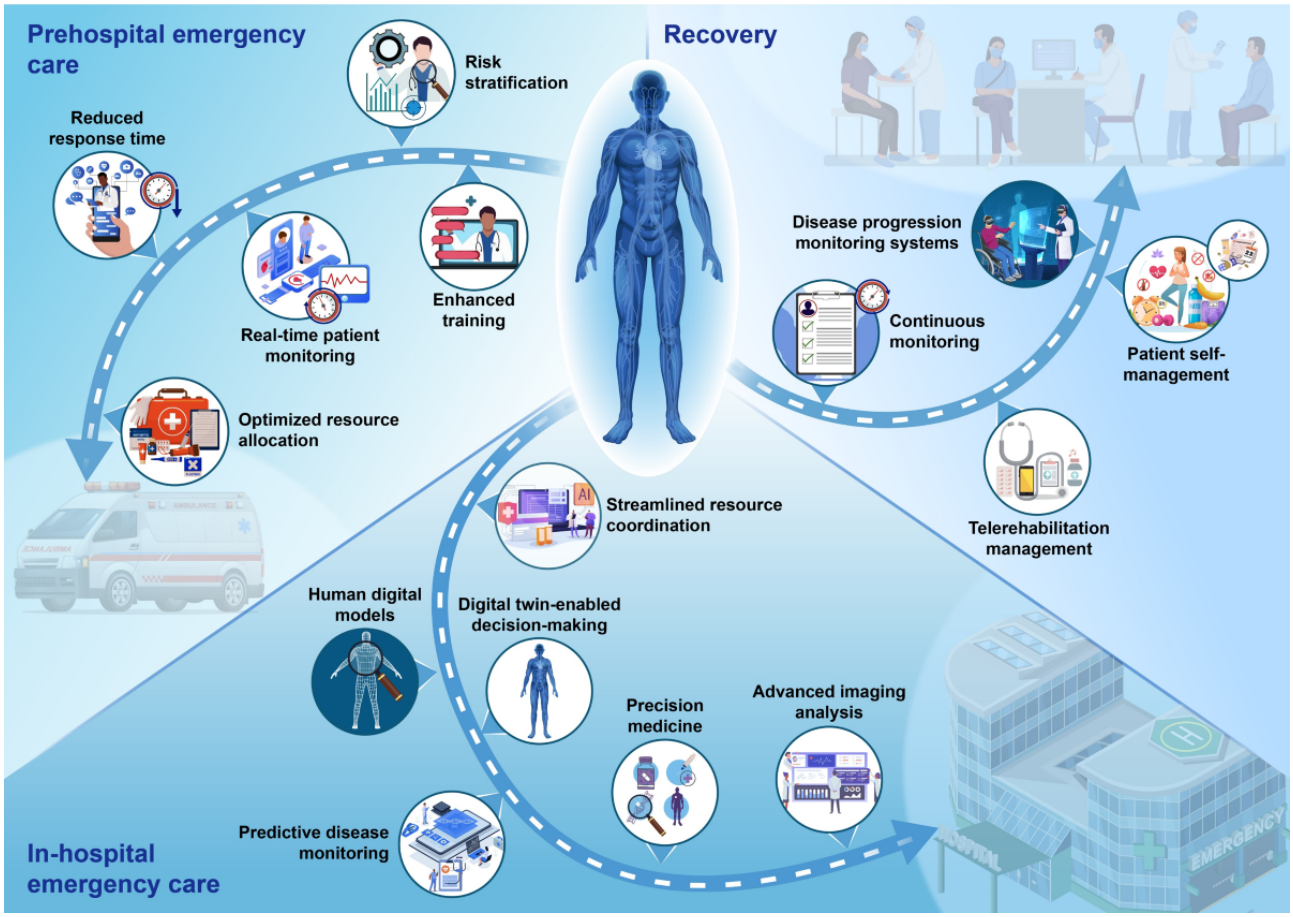
KEYWORDS

emergency care; digital twin; prehospital emergency care; in-hospital emergency care; recovery

Introduction

The concept of digital twins—dynamic, real-time simulations of processes, systems, or environments—has garnered increasing attention across various domains, from urban planning to medical monitoring [1]. Unlike predictive analytics and artificial intelligence (AI)–driven simulations, digital twins construct dynamic mirror models of physical entities through real-time data synchronization and simulation technologies, emphasizing virtual to physical mapping and closed-loop feedback to support full lifecycle management. In contrast, predictive analytics focus on trend forecasting based on historical data, whereas AI-driven simulations concentrate on multiscenario modeling of complex systems. The advantages of digital twins lie in their real-time capabilities, dynamic optimization, and potential for cross-system collaboration, enabling a closed-loop process from monitoring to decision-making, significantly enhancing efficiency and precision [2]. Emergency care, characterized by its demand for timeliness, precision, and adaptability, provides an ideal testing ground for this transformative technology. By bridging prehospital care, in-hospital management, and recovery, digital twins promise to reshape emergency medical systems for the future (Figure 1). Textbox 1 provides an explanation of the application of digital twins in prehospital first aid, in-hospital first aid, and rehabilitation.

Figure 1. The role of digital twins in transforming emergency medicine.



Textbox 1. The application of digital twins in prehospital first aid, in-hospital first aid, and rehabilitation.

Prehospital emergency care

- Optimized resource allocation: simulation of real-time first-aid scenarios to assess urban emergency needs and spatial distribution of resources, allowing for evidence-based planning and improved efficiency in resource use
- Reduced response time: minimization of procedural delays through real-time data sharing, enabling seamless communication and critical data updates across emergency personnel and systems
- Real-time patient monitoring: leveraging of wearable devices and sensors to create dynamic health profiles, facilitating immediate assessments and predictions of patient conditions (eg, vital signs such as heart rate and blood pressure)
- Risk stratification: aggregation and analysis of patient data to provide superior accuracy in predicting acute events (eg, cardiac arrests), outperforming traditional clinical risk assessments
- Enhanced training: use of simulation-based training environments to strengthen emergency response protocols, improve decision-making, and foster team coordination under high-pressure conditions

In-hospital emergency care

- Advanced imaging analysis: integration of multidimensional data to overcome sample limitations in medical imaging, enabling the development of robust diagnostic models
- Precision medicine: combination of genomic, environmental, and lifestyle data for tailored risk assessments and personalized treatment plans, empowering informed clinical decision-making
- Human digital models: use of digital replicas of anatomical structures to plan surgeries and minimize risks during procedures by simulating operations beforehand
- Digital twin–enabled decision-making: facilitation of real-time analysis and expert-guided interventions (eg, cardiopulmonary resuscitation and drug administration) via remote systems, reducing disparities in emergency care accessibility
- Predictive disease monitoring: use of Internet of Things devices and wearables to continuously monitor patient status, identifying early signs of complications and supporting timely interventions
- Streamlined resource coordination: ensuring efficient resource distribution and effective communication among health care teams through unified, up-to-date patient data sharing

Recovery

- Continuous monitoring: use of sensors and wearable devices to track recovery progress and dynamically adjust treatment plans to optimize outcomes
- Telerehabilitation management: enabling remote monitoring of patients through digital twins, providing clinicians with tools to assess and adjust rehabilitation remotely, particularly for mobility-impaired individuals
- Patient self-management: sharing digital twin insights with patients to encourage active participation in health management through lifestyle modifications and adherence to care plans
- Disease progression monitoring systems: integration of real-time monitoring data to predict the risk of disease exacerbations or complications, thereby intervening in time and triggering early warnings, allowing physicians to adjust treatment during critical windows and improve outcomes.

Prehospital Emergency Care

Current prehospital emergency care systems often struggle with suboptimal resource allocation due to their reliance on simplified decision-making models that fail to account for the complex, dynamic nature of urban environments [3]. Digital twins address these limitations by unifying diverse data streams into a single, adaptable model that enables real-time optimization of emergency resource allocation. These models incorporate multiple layers of data, from historical emergency call patterns to real-time traffic conditions, enabling more nuanced and effective resource deployment decisions.

Digital twins prove particularly valuable in prehospital emergency care, where time-sensitive decisions directly impact patient outcomes. By incorporating high-resolution information about building layouts, land use, population density, and traffic patterns, digital twins create actionable urban simulations that allow emergency response teams to plan and execute their operations with unprecedented efficiency. The integration of these various data sources enables more sophisticated approaches to emergency resource management than previously possible with traditional systems.

This transformation in emergency care delivery demonstrates the broader potential of digital twin technology to revolutionize urban health care systems through data-driven, predictive approaches to resource management. Moreover, wearables in ambulances provide dynamic health profiles, supporting real-time monitoring and accurate risk stratification for acute events. Simulation-based training environments can help improve emergency team readiness and coordination under high-pressure conditions.

In-Hospital Emergency Care

In hospitals, digital twins are redefining precision medicine by integrating diverse data sources to enable robust diagnostics, personalized treatments, and streamlined workflows [4]. By synthesizing real-time patient data with historical health records and predictive algorithms, digital twins facilitate more accurate disease modeling, enhancing clinical decision-making and tailoring interventions to individual patient profiles. This data-driven approach supports the early detection of complications, optimization of treatment pathways, and improvement of patient outcomes.

Preoperative simulations represent a critical application of digital twins in surgical care. By creating highly accurate digital replicas of patient anatomy, clinicians can simulate surgical procedures, anticipate potential challenges, and refine techniques before making the first incision. These virtual rehearsals significantly reduce surgical risks, improve procedural efficiency, and enhance patient safety. Complementing this, predictive monitoring through Internet of Things (IoT) devices enables the continuous assessment of patient vital signs, offering early warnings of complications such as sepsis or cardiac arrest. These alerts allow medical teams to respond swiftly, averting adverse outcomes and saving lives.

An emerging and transformative application is the development of medical device twins (MDTs), which provide virtual counterparts for critical hospital equipment, including ventilators, defibrillators, infusion pumps, and imaging devices. MDTs are created through digital twin technology and are consistent with the actual medical device in terms of function, performance, and status. They can be updated in real time to reflect the actual operation of the medical device. These digital twins continuously monitor device performance, tracking metrics such as wear and tear, calibration accuracy, and operational efficiency. By identifying potential failures before they occur, MDTs facilitate proactive maintenance, reducing downtime and ensuring the reliability of life-saving equipment during critical emergencies.

Beyond maintenance, MDTs contribute to resource optimization. By analyzing use patterns and forecasting demand surges, these systems can dynamically allocate equipment to areas of highest need, such as during mass casualty incidents or seasonal patient influxes. This predictive capability not only prevents bottlenecks in care delivery but also enhances hospital operational efficiency, ensuring that critical resources are always available when and where they are needed.

Recovery

In recovery, digital twins play a transformative role by enabling personalized and adaptive rehabilitation processes. These advanced systems integrate comprehensive patient data, including clinical records, biometric parameters, and lifestyle information, to develop dynamic treatment plans tailored to individual needs. Wearable devices continuously monitor patient progress in real time, providing crucial data to adjust therapies dynamically, ensuring that interventions remain effective and responsive to changes in a patient's condition [5,6]. This real-time feedback loop facilitates precision rehabilitation, minimizing setbacks and promoting steady recovery trajectories.

Telerehabilitation further amplifies these benefits by offering remote monitoring and therapy-adjustment capabilities. Patients can receive professional guidance and care modification without frequent in-person visits, reducing logistical burdens and enhancing access to rehabilitation services. Moreover, digital twins empower patients by offering insights into their recovery progress, encouraging active engagement in health management and fostering adherence to prescribed care plans.

A significant breakthrough in the recovery phase is disease relapse prediction and early warning, a feature that underscores the predictive power of digital twins. By leveraging longitudinal patient data and advanced predictive analytics, digital twin systems can detect subtle patterns indicative of potential disease recurrence. These insights enable the generation of real-time alerts for health care providers, prompting timely interventions that can mitigate complications and prevent hospital readmissions. For instance, changes in respiratory patterns or biomarkers could signal the early stages of a relapse, allowing for preemptive adjustments in treatment strategies.

Digital Twins in Real-World Applications

Currently, the application of digital twins in emergency care remains in its early developmental stages, with significant room for growth, yet they have already demonstrated potential in real-world scenarios. For instance, Pfizer and IBM's collaborative Project BlueSky focuses on developing remote patient monitoring systems that use wearable devices and environmental sensors to track disease symptoms in real time, capture dynamic data, and automatically identify abnormalities for timely intervention [7]. Similarly, Digital Orthopaedics and Dassault Systèmes jointly created a clinical decision support system capable of generating patient-specific computational models to reduce trial-and-error risks and achieve precision treatment [7]. During public health crises like COVID-19, digital twins have proven effective in enhancing emergency response capabilities by simulating real-time data collection and vaccination processes to identify optimal resource allocation strategies and refine contingency planning. These examples underscore the transformative role of digital twin technology in advancing emergency care efficiency and resilience [8].

In the cases mentioned above, digital twins have demonstrated significant advantages over current public health systems. Digital public health enables optimized resource allocation, reduces waste, and enhances emergency response efficiency, and it allows real-time monitoring of patient anomalies to accelerate emergency interventions. Centered on data-driven and intelligent capabilities, these advancements markedly improve the efficiency of resource allocation, response speed, and decision-making precision in emergency care.

In summary, digital twins represent a reliable tool in the era of digital intelligence, becoming increasingly feasible in the current data-rich context. The emergency response system serves as an excellent testing ground for digital twin applications, offering

significant potential across areas such as emergency resource planning, patient monitoring, and rehabilitation. However, the implementation of digital twins in the emergency domain is not a task to be completed overnight. It necessitates advancements in complementary fields, including urban modeling, wearable technology, and the IoT, to fully realize its potential. Despite their promise, the application of digital twins in emergency care faces substantial challenges. Currently, many medical digital twins remain confined to visualization, providing static representations rather than actionable tools for decision-making. For full realization, these systems must evolve beyond visualization to actionable frameworks. A second critical hurdle is data quality and resolution. Biological heterogeneity significantly influences disease progression and treatment outcomes, requiring vast, high-resolution datasets to build robust, accurate models. Lastly, ethical, privacy, and security concerns pose significant barriers to the adoption of digital twins in health care [9]. The integration of sensitive patient data with urban-level modeling raises important questions about data governance, requiring careful consideration to balance innovation with public trust. To address the aforementioned challenges, it is necessary to systematically overcome existing bottlenecks through technological optimization, privacy protection, ethical governance, and multiparty collaboration. On the privacy and security front, full lifecycle data protection must be strengthened by leveraging differential privacy and blockchain technology to anonymize sensitive information and enable traceable storage while implementing dynamic access controls to restrict unnecessary data access. Ethically and legally, clear boundaries and transparent guidelines for technology use must be established. Additionally, collaboration among governments, medical institutions, and technology companies is needed to develop unified technical standards and security evaluation systems alongside public education to build societal trust and ensure patient consent and data authorization. Only through parallel advancements in technology, regulatory frameworks, ethical constraints, and social collaboration can the safe implementation and sustainable development of digital twin technology in emergency scenarios be achieved.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (72404117). The funder had no role in the design or conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Authors' Contributions

BZ and HL designed and supervised the research. HL drafted the manuscript. BZ revised the manuscript. JZ and NZ commented on the revision of the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Chen J, Yi C, Okegbile SD, Cai J, Shen X. Networking architecture and key supporting technologies for human digital twin in personalized healthcare: a comprehensive survey. IEEE Commun Surv Tut 2024;26(1):706-746. [doi: 10.1109/comst.2023.3308717]
2. Li X, Feng M, Ran Y, Su Y, Liu F, Huang C, et al. Big data in earth system science and progress towards a digital twin. Nat Rev Earth Environ 2023 May 02;4(5):319-332. [doi: 10.1038/s43017-023-00409-w]
3. Fan C, Zhang C, Yahja A, Mostafavi A. Disaster city digital twin: a vision for integrating artificial and human intelligence for disaster management. Int J Inf Manage 2021 Feb;56:102049. [doi: 10.1016/j.ijinfomgt.2019.102049]
4. Laubenbacher R, Mehrad B, Shmulevich I, Trayanova N. Digital twins in medicine. Nat Comput Sci 2024 Mar;4(3):184-191. [doi: 10.1038/s43588-024-00607-6] [Medline: 38532133]
5. Sun T, He X, Li Z. Digital twin in healthcare: recent updates and challenges. Digit Health 2023;9:20552076221149651 [FREE Full text] [doi: 10.1177/20552076221149651] [Medline: 36636729]
6. Scott AK, Oyen ML. Virtual pregnancies: predicting and preventing pregnancy complications with digital twins. Lancet Digit Health 2024 Jul;6(7):e436-e437 [FREE Full text] [doi: 10.1016/S2589-7500(24)00086-4] [Medline: 38906606]
7. Mohapatra S, Bose S. An appraisal of literature for design and implementation of developing a framework for digital twin and validation through case studies. Health Technol 2020 Jul 04;10(5):1229-1237. [doi: 10.1007/s12553-020-00443-4]
8. De Benedictis A, Mazzocca N, Somma A, Strigaro C. Digital twins in healthcare: an architectural proposal and its application in a social distancing case study. IEEE J Biomed Health Inform 2023 Oct;27(10):5143-5154. [doi: 10.1109/JBHI.2022.3205506] [Medline: 36083955]
9. Willcox K, Segundo B. The role of computational science in digital twins. Nat Comput Sci 2024 Mar;4(3):147-149. [doi: 10.1038/s43588-024-00609-4] [Medline: 38532131]

Abbreviations

AI: artificial intelligence
IoT: Internet of Things
MDT: medical device twin

Edited by Y Luo; submitted 26.01.25; peer-reviewed by AK Jameil, J Wang; comments to author 10.03.25; revised version received 22.03.25; accepted 12.04.25; published 21.04.25.

Please cite as:
Li H, Zhang J, Zhang N, Zhu B
Advancing Emergency Care With Digital Twins
JMIR Aging 2025;8:e71777
URL: https://aging.jmir.org/2025/1/e71777
doi:10.2196/71777
PMID:

Machine Learning for Predicting Postoperative Functional Disability and Mortality Among Older Patients With Cancer: Retrospective Cohort Study

Yuki Hashimoto[1], PhD; Norihiko Inoue[1], MD, PhD; Takuaki Tani[2], PhD; Shinobu Imai[2], PhD

[1]Department of Clinical Data Management and Research, Clinical Research Center, National Hospital Organization Headquarters, 2-5-21 Higashigaoka, Meguroku, Japan

[2]Department of Pharmacoepidemiology, Showa University Graduate School of Pharmacy, Shinagawaku, Japan

Corresponding Author:
Yuki Hashimoto, PhD
Department of Clinical Data Management and Research, Clinical Research Center, National Hospital Organization Headquarters, 2-5-21 Higashigaoka, Meguroku, Japan

Abstract

Background: The global cancer burden is rapidly increasing, with 20 million new cases estimated in 2022. The world population aged ≥65 years is also increasing, projected to reach 15.9% by 2050, making cancer control for older patients urgent. Surgical resection is important for cancer treatment; however, predicting postoperative disability and mortality in older patients is crucial for surgical decision-making, considering the quality of life and care burden. Currently, no model directly predicts postoperative functional disability in this population.

Objective: We aimed to develop and validate machine-learning models to predict postoperative functional disability (≥5-point decrease in the Barthel Index) or in-hospital death in patients with cancer aged ≥ 65 years.

Methods: This retrospective cohort study included patients aged ≥65 years who underwent surgery for major cancers (lung, stomach, colorectal, liver, pancreatic, breast, or prostate cancer) between April 2016 and March 2023 in 70 Japanese hospitals across 6 regional groups. One group was randomly selected for external validation, while the remaining 5 groups were randomly divided into training (70%) and internal validation (30%) sets. Predictor variables were selected from 37 routinely available preoperative factors through electronic medical records (age, sex, income, comorbidities, laboratory values, and vital signs) using crude odds ratios (P<.1) and the least absolute shrinkage and selection operator method. We developed 6 machine-learning models, including category boosting (CatBoost), extreme gradient boosting (XGBoost), logistic regression, neural networks, random forest, and support vector machine. Model predictive performance was evaluated using the area under the receiver operating characteristic curve (AUC) with 95% CI. We used the Shapley additive explanations (SHAP) method to evaluate contribution to the predictive performance for each predictor variable.

Results: This study included 33,355 patients in the training, 14,294 in the internal validation, and 6711 in the external validation sets. In the training set, 1406/33,355 (4.2%) patients experienced worse discharge. A total of 24 predictor variables were selected for the final models. CatBoost and XGBoost achieved the largest AUCs among the 6 models: 0.81 (95% CI 0.80-0.82) and 0.81 (95% CI 0.80-0.82), respectively. In the top 15 influential factors based on the mean absolute SHAP value, both models shared the same 14 factors such as dementia, age ≥85 years, and gastrointestinal cancer. The CatBoost model showed the largest AUCs in both internal (0.77, 95% CI 0.75-0.79) and external validation (0.72, 95% CI 0.68-0.75).

Conclusions: The CatBoost model demonstrated good performance in predicting postoperative outcomes for older patients with cancer using routinely available preoperative factors. The robustness of these findings was supported by the identical top influential factors between the CatBoost and XGBoost models. This model could support surgical decision-making while considering postoperative quality of life and care burden, with potential for implementation through electronic health records.

(JMIR Aging 2025;8:e65898) doi:10.2196/65898

KEYWORDS

older patients with cancer; postoperative outcomes; functional disability; machine learning; decision-making

Introduction

The global cancer burden is rapidly increasing, with an estimated 20 million new cases and 9.7 million deaths in 2022 [1]. In Japan, the lifetime risk of being diagnosed with cancer is approximately 65.5% and 51.2% for men and women, respectively [2].

In addition, the global population is aging rapidly, with the proportion of those aged ≥65 years expected to increase from 9.1% in 2019 to 15.9% by 2050 [3]. Japan faces the most advanced stage of this demographic shift, with the population of older adults expected to increase from 28.8% in 2020 to 37.7% by 2050 [4]. Thus, cancer control for older patients has become an urgent issue worldwide, including in Japan.

Older patients with cancer often face challenges such as frailty [5], comorbidities [6-8], and socioeconomic status [6], which affect treatment outcomes and quality of life (QOL). Surgical resection is a key treatment, which requires careful consideration in older patients with cancer due to concerns about postoperative functional disability and its impact on long-term outcomes [9]. Some factors may influence surgical outcomes in older patients, including age [9], anemia [10], BMI [11,12], dementia [6,7], frailty [5], low household income [6], malnutrition [13], and smoking [14].

Considering the postoperative QOL and care burden on patients' families and society, it is important to predict not only postoperative mortality but also functional disability [15]. Hospital-associated disability, defined as functional disability following acute hospitalization, is recognized as a crucial outcome in older patients with significant impact on health care costs and long-term prognosis [16,17]. Some models can predict postoperative mortality [5,13]; however, few have addressed functional disability. A model after lower-extremity surgery [6] could predict the risk of in-hospital mortality and discharge to a nursing home, which is a surrogate for functional disability, in patients admitted from home. Currently, no model can directly predict functional disability after cancer surgery.

Therefore, this study aimed to develop a machine learning–based model for predicting postoperative functional disability and in-hospital mortality in patients with cancer aged ≥65 years using data from 70 hospitals across Japan. This approach will enable patients and their families to make informed decisions about undergoing cancer surgery, considering their postoperative QOL and care burden.

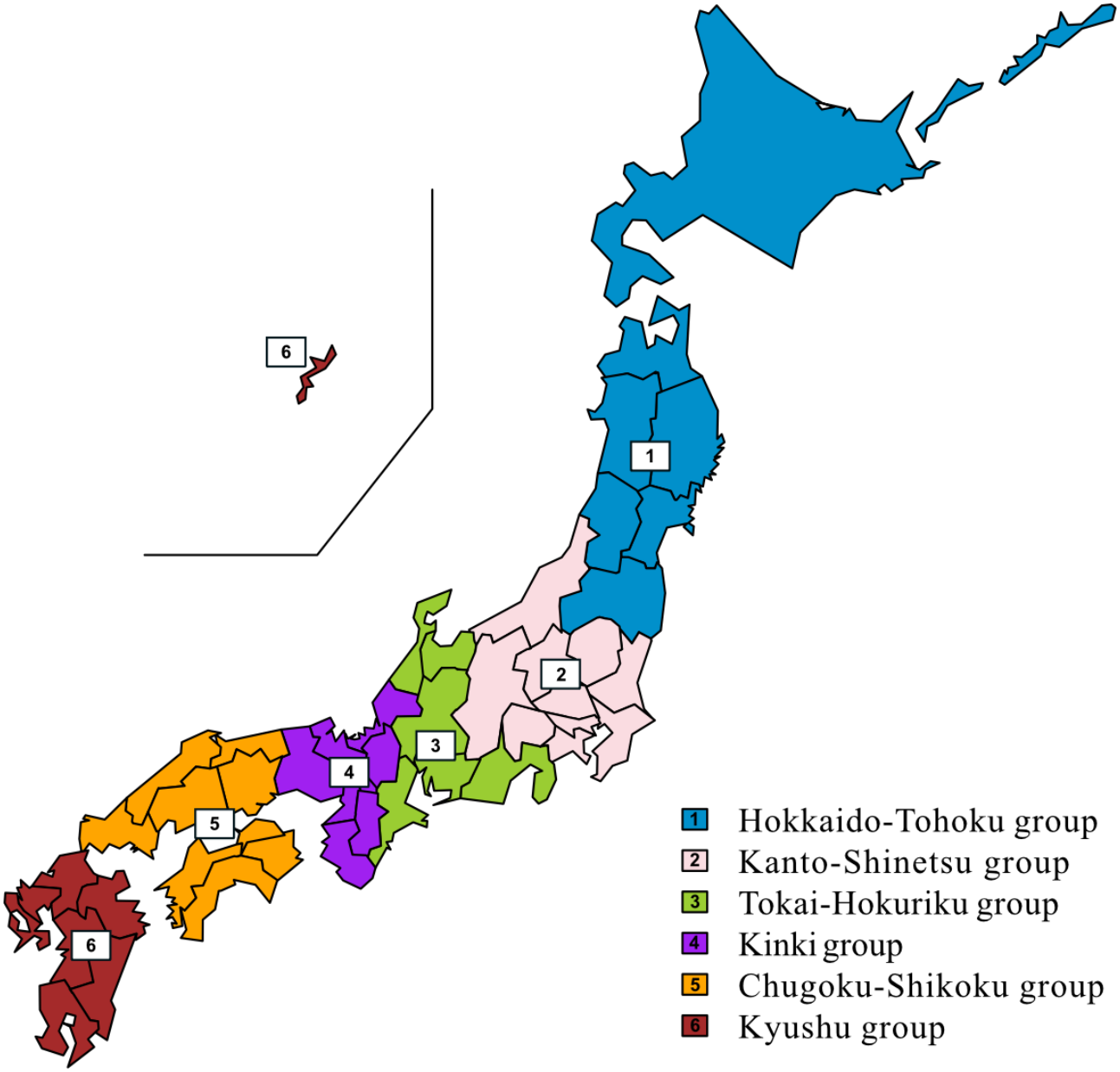
Methods

Study Design and Data Source

We conducted a retrospective cohort study to develop machine-learning models for predicting postoperative functional disability and in-hospital mortality in older patients with cancer. This study used data between April 2016 and March 2023 from 70 Japanese hospitals within the National Hospital Organization (NHO) database across 6 regional groups: Hokkaido-Tohoku group, Kanto-Shinetsu group, Tokai-Hokuriku group, Kinki group, Chugoku-Shikoku group, and Kyushu group (Figure 1) [18].

The NHO maintains 2 databases: (1) an administrative claims database based on the Diagnosis Procedure Combination–based Per-Diem Payment System [19] and a clinical information database based on the standardized structured medical record information exchange [20]. The administrative claims database contains patient information, such as age, sex, cost, comorbidities, complications, diagnosis, medical procedures, and medications. The clinical information database includes medical charts, laboratory data, and vital signs on a daily basis.

Figure 1. A total of 6 National Hospital Organization regional groups across Japan.



Participants

The study included patients aged ≥65 years who were admitted to NHO hospitals between April 2016 and March 2023 and underwent surgery for major cancers, including lung, stomach, colorectal, liver, pancreatic, breast, and prostate cancers. These cancer sites were selected because of their high incidence [2] and mortality rates [21] in Japanese and global cancer statistics [1]. The surgical procedures included both scopic and open surgeries under general anesthesia.

We excluded patients who had missing Barthel Index data at admission or discharge, were first included in the database at admission (no medical history available), underwent surgery more than 1 week after admission, or had the Barthel Index of 0 at admission. These exclusion criteria were applied because: (1) patients missing Barthel Index data could not be evaluated for outcome, (2) most of the predictor variables were missing if patients had no medical history, (3) we eliminated the effect of hospitalization on physical function from admission to surgery, and (4) the Barthel Index change from admission to discharge was an outcome variable, but patients with a minimum score of 0 at admission cannot show further decline [22].

Outcome Variables

The primary outcome was worse discharge, defined as either in-hospital death or postoperative functional disability. Postoperative functional disability was characterized as hospital-associated disability [16] (≥5-point decrease in the Barthel Index between admission and discharge) [17]. The Barthel Index consists of 10 items, including transfer, bathing, and stair climbing, used to evaluate activities of daily living on a scale of 0 - 100, with lower scores indicating a decline in physical function.

The secondary outcomes were health care costs and postoperative length of stay (LOS) in patients predicted to be at high risk and low risk for worse discharge. The optimal cut-off point for high-risk or low-risk classification was

determined using the Youden index [23] on the receiver operating characteristic (ROC) curve.

Predictor Variables

We identified the following 37 potential predictors of worse discharge based on previous studies: age (65 - 74, 75 - 84, and ≥85 years) [9], sex [24], underweight (BMI <18.5 kg/m^2) [12] and obesity (BMI ≥30 kg/m^2) [11], route of admission (home or nonhome), emergency admission, an estimated household income in the lowest tertile based on post-code (low income) [6], smoking (Brinkman Index ≥200) [14], functional dependence (Barthel Index ≤60) [22], surgical factors (open surgery, scopic surgery, and combined general and epidural anesthesia), the presence of gastrointestinal cancer (colorectal, liver, pancreas, and stomach), the cancer staging (0–II or III–IV) [25], the presence of recurrent cancer, Charlson Comorbidity Index (CCI) ≥3 [26,27], Hospital Frailty Risk Score (HFRS) ≥5 [28], comorbidities common in older patients (cerebrovascular disease, chronic pulmonary disease, congestive heart failure, dementia [8], diabetes, liver disease, myocardial infarction, peptic ulcer disease, peripheral vascular disease, and renal disease), medical history about cancer within 8 weeks before surgery that allowed the evaluation of preoperative chemotherapy [29] and radiotherapy [30] (chemotherapy, radiation, and surgery), vital signs (body temperature ≥38°C [31], systolic blood pressure >180 mm Hg [32]), and laboratory test values (albumin <3.5 g/dL [13], total bilirubin ≥2.0 mg/dL [33], creatinine ≥ 2.0 mg/dL [33], platelet <10^5/μL [33], and hemoglobin <11 g/dL [10]).

These predictors were measured as follows: age, sex, BMI, route of admission, emergency admission, income, cancer staging evaluated using the Tumor Nodes Metastasis (TNM) classification system [34], the presence of recurrent cancer, Brinkman Index, and Barthel Index were assessed at admission. Surgical factors were assessed during the surgery. The type of cancer was determined during the surgery using the *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* (*ICD-10*) coding system [35]. Comorbidities were identified from *ICD-10* codes within 8 weeks before surgery (Table S1 in Multimedia Appendix 1). Notably, dementia was determined if it was either identified from *ICD-10* codes within 8 weeks before surgery or documented in the clinical summary at admission, as Japanese hospitals are required to include the dementia status of inpatients aged ≥65 years at admission [8]. Vital signs were measured closest to surgery after admission, and laboratory values were obtained using those measured closest to surgery within 8 weeks before surgery.

Statistical Analysis

We randomly selected one of the 6 NHO regional groups for the external validation set [36], while hospital data from the remaining 5 groups were randomly divided into the training (70%) and the internal validation (30%) sets [37]. The missing predictor variables were imputed using the "missRanger" [38], which is a random forest–based algorithm [39], assuming that the data are missing at random.

To summarize patient characteristics, continuous variables were expressed as mean (SD) or median (IQR), depending on the distribution of variables. The Wilcoxon rank-sum test or the Welch test was used to assess between-group differences. Categorical variables were expressed as proportions and compared using the χ^2 test.

In the training set, a double penalty was implemented by eliminating unnecessary variables to create a more practical model for clinical use [40]. The first penalty involved selecting predictor candidates with a crude odds ratio (OR) at P<.1 [41]. The second penalty to further narrow down the predictor candidates used the least absolute shrinkage and selection operator (Lasso) method, which allowed for the selection of clinically relevant variables with consistent relationships [36,42].

The selected factors were incorporated into 6 machine learning models: category boosting (CatBoost) [36,43], extreme gradient boosting (XGBoost) [36,42], logistic regression, neural networks [44], random forest [36], and support vector machine (SVM) [36]. Model performance was evaluated using the area under the ROC curve (AUC) with 95% CIs). Similarly, we calculated the accuracy, sensitivity, specificity, F_1-score, and the area under the precision-recall curve (PRAUC) to assess the performance of the models [36,37]. The precision-recall curve of the models was also shown.

We used the synthetic minority oversampling and random undersampling techniques to avoid overfitting owing to the imbalance between the positive and negative events [45]. The minority class was oversampled at 50%, 100%, and 200% of its original size, followed by random undersampling of the majority class to achieve equal numbers between classes. The models were trained using 10-fold cross-validation with grid search for hyperparameter optimization. Of these sampling ratios and hyperparameter combinations, those yielding the largest AUC were selected. We ranked the predictor variable using the Shapley additive explanations (SHAP) method [46] to assess the contribution of the predictors to the models. Moreover, considering higher cancer stages are associated with poorer postoperative outcomes [25], a multiple logistic regression was conducted to evaluate the interaction between cancer staging and other predictor variables using OR (95% CI). We examined interactions between cancer stage III–IV and the top 5 features based on the mean absolute SHAP value.

We analyzed the AUC difference between the models using the DeLong method [47], which was based on the model with the largest AUC in both internal and external validation. Sensitivity analyses were used to assess the impact of missing data and assessing 2 outcomes simultaneously (death or functional disability) on the models. For missing values, a complete case analysis was performed to confirm the robustness of the results obtained from the imputed data set. We also evaluated the AUC of the models when the models predicted only death in all patients or functional disability in patients with survival discharge. We conducted subgroup analyses by LOS, cancer type (breast, colorectal, liver, lung, pancreas, prostate, or stomach), and cancer staging (stage 0–I, II, III, or IV). For the LOS analysis, we calculated the 75th percentile of LOS for each cancer type separately and divided patients into LOS for each

cancer type <75th percentile (short-stay) and LOS ≥75th percentile (long-stay) groups [48].

For the analysis of secondary outcomes, we compared LOS and health care costs between patients at high and low risk based on the model with the highest AUC in both internal and external validation. Health care costs were assessed across various categories, including total, medical consultation, medication, medical procedure, surgical procedure, laboratory tests, hospital stay, and others. The currency conversion rate was 150 JPY to US $1.

All hypothesis tests had a 2-sided significance level of .05. All statistical analyses were performed using R version 4.3.1 (R Foundation for Statistical Computing) .

Ethical Considerations

Our study was approved by the Institutional Review Board of Showa University (approval number 2023 - 129-A). Individual consent was not required because this was an opt-out study. This study conforms to the principles outlined in the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.

Results

In total, 54,360 patients from 70 hospitals were included in this study (Figure 2): 6711 in the external validation set from Kinki group, 33,355 in the training set, and 14,294 in the internal validation set from the remaining 5 regional groups. In the training set, 1406/33,355 (4.2%) patients experienced worse discharge (Table 1). These patients were older (24.5% [344/1406] vs 5.8% [1864/31,949] for age ≥85 years), more likely to be admitted from nonhome (8.1% [114/1406] vs 0.9% [281/31,949]), and had higher prevalence of dementia (28.7% [403/1406] vs 5.6% [1783/31,949]), gastrointestinal cancer (64.3% [904/1406] vs 41.7% [13,326/31,949]), and advanced cancer stage (32.6% [458/1406] vs 24.2% [7716/31,9] for stage III–IV).

In the training set, we selected 31 predictor variables for worse discharge using crude OR (Table S2 in Multimedia Appendix 1). Further selection using the Lasso method resulted in 24 factors: age (75 - 84 y and ≥85 y), male sex, BMI <18.5 kg/m^2, nonhome admission, emergency admission, low income, Barthel Index at admission ≤60, open surgery, gastrointestinal cancer, cancer stage III–IV, HFRS ≥5, cerebrovascular disease, congestive heart failure, dementia, diabetes, liver disease, myocardial infarction, medical history of chemotherapy, systolic blood pressure ≥180 mm Hg, albumin<3.5 g/dL, creatinine ≥ 2.0 mg/dL, platelet <10^5/μL, and hemoglobin <11 g/dL. All 6 models were developed using these 24 variables.

Furthermore, the AUCs were 0.81 (95% CI 0.80 - 0.82) for CatBoost, 0.81 (95% CI 0.80 - 0.82) for XGBoost, 0.79 (95% CI 0.78 - 0.80) for random forest, 0.79 (95% CI 0.78 - 0.80) for neural networks, 0.78 (95% CI 0.77 - 0.80) for SVM, and 0.78 (95% CI 0.77 - 0.80) for logistic regression (Figure 3). CatBoost and XGBoost were the 2 models with AUC ≥0.80 and showed similar values with relatively high accuracy (0.76), sensitivity (0.72), specificity (0.76), F_1-score (0.20), and area under the precision-recall curve (PRAUC) (0.22). The performance metrics and precision-recall curves for all models are shown in Table S3 and Figure S1 in Multimedia Appendix 1, respectively. In the top 15 influential factors based on the mean absolute SHAP value, the CatBoost and XGBoost models had the same combination for the 14 features: dementia, age ≥85 years, age 74 - 85 years, gastrointestinal cancer, albumin <3.5 g/dL, open surgery, male sex, hemoglobin <11 g/dL, low income, nonhome admission, Barthel Index at admission ≤60, BMI <18.5 kg/m^2, diabetes, and stage III–IV (Figure 4). There were no significant interactions between stage III–IV and the top 5 influential features that both models shared (Table S4) in Multimedia Appendix 1.

For both the internal and external validation set, the CatBoost model had the largest AUCs among the 6 machine- learning models: 0.77 (95% CI 0.75 - 0.79) and 0.72 (95% CI 0.68 - 0.75), respectively (Figure 5). In sensitivity analysis, all 6 models maintained comparable performance to the main analysis (Table S5 in Multimedia Appendix 1). The CatBoost model achieved relatively high AUCs and showed consistent performance in complete cases (0.78, 95% CI 0.76 - 0.80; 0.72, 0.68 - 0.76), death only (0.77, 0.71 - 0.82; 0.73, 0.65 - 0.81), and functional disability only (0.77, 0.75 - 0.79; 0.71, 0.68 - 0.75) for internal and external validation, respectively.

In subgroup analyses, the models maintained consistent performance for LOS and cancer staging (Table S5 in Multimedia Appendix 1). However, the performance varied based on cancer types: the CatBoost model achieved a larger AUC in patients with stomach cancer (internal: 0.80, 95% CI 0.76 - 0.84; external: 0.81, 95% CI 0.69 - 0.92) but a smaller AUC in patients with prostate cancer (0.53, 95% CI 0.40 - 0.66; 0.46, 95% CI 0.28 - 0.63) than the main analysis.

Based on the CatBoost model, patients at high risk had significantly longer LOS (internal: median 13, IQR 9 - 19 d vs median 9, IQR 7 - 13 d; external: median 13, IQR 10 - 19 d vs median 10.0, IQR 7.0 - 14.0 d) and higher total health care costs (internal: median US $11,048, IQR US $9191 - 13,106 d vs median US $10,092, IQR US $7894 - 11,893; external: median 11,069, IQR US $9401 - 13,499 vs median US $10,371, IQR US $8820 - 11,936) than patients at low risk (all P<.01). However, the high-risk group had slightly lower surgical procedure costs than the low-risk group in internal validation and was comparable to the low-risk group in external validation (Table 2).

Figure 2. Flow diagram of enrollment of study participants.(a) “Missing BI” means patients with missing Barthel Index at admission or discharge. (b) “No medical history” means patients included in the database within 8 weeks preceding surgery. (c) “More than 1 week” means patients who underwent surgery more than one week after admission. (d) “BI at admission=0” means patients with BI of 0 at admission. BI: Barthel Index; NHO: National Hospital Organization.

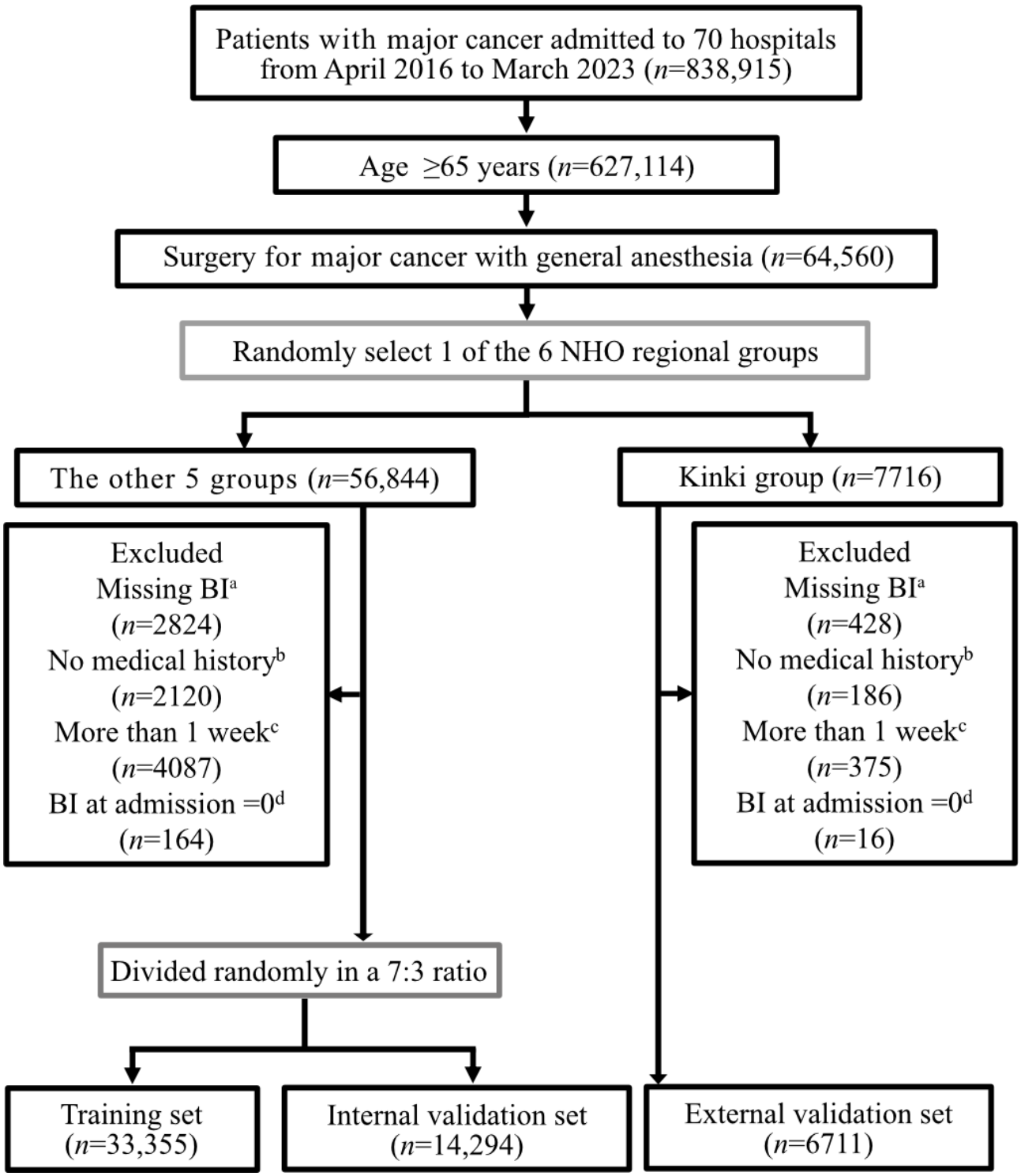


Table . Patient background with or without worse discharge in the training set.

Variable	No worse discharge (n=31,949)	Worse discharge (n=1406)	*P* value
Age (years), n (%)			<.01
65 - 74	18,010 (56.4)	405 (28.8)	
75 - 84	12,075 (37.8)	657 (46.7)	
≥85	1864 (5.8)	344 (24.5)	
Sex, n (%)			.05
Male	16,570 (51.9)	768 (54.6)	
Female	15,379 (48.1)	638 (45.4)	
BMI (kg/m^2), n (%)			
<18.5	2535 (7.9)	212 (15.1)	<.01
≥30	1170 (3.7)	57 (4.1)	.49
Route of admission, n (%)			
Home	31,668 (99.1)	1292 (91.9)	<.01
Nonhome	281 (0.9)	114 (8.1)	<.01
Nursing home	134 (0.4)	60 (4.3)	<.01
Other hospital	136 (0.4)	54 (3.8)	<.01
Others	11 (0)	0 (0)	1.00
Emergency admission, n (%)	252 (0.8)	42 (3)	<.01
Low income[a], n (%)	6020 (18.8)	352 (25)	<.01
Brinkman Index ≥200, n (%)	13,307 (41.7)	551 (39.2)	.07
Barthel Index[b]≤60, n (%)	422 (1.3)	139 (9.9)	<.01
Open surgery, n (%)	12,307 (38.5)	652 (46.4)	<.01
Scopic surgery, n (%)	19,642 (61.5)	754 (53.6)	.01
With epidural anesthesia[c], n (%)	15,742 (49.3)	769 (54.7)	<.01
Type of cancer, n (%)			<.01
Breast	6298 (19.7)	150 (10.7)	
Colorectal	7387 (23.1)	493 (35.1)	
Liver	1331 (4.2)	91 (6.5)	
Lung	9386 (29.4)	293 (20.8)	
Pancreas	786 (2.5)	62 (4.4)	
Prostate	2939 (9.2)	59 (4.2)	
Stomach	3822 (12)	258 (18.3)	
Gastrointestinal cancer	13,326 (41.7)	904 (64.3)	<.01
Cancer staging, n (%)			<.01
0-I	14,398 (45.1)	492 (35)	
II	9835 (30.8)	456 (32.4)	
III	4845 (15.2)	316 (22.5)	
IV	2871 (9)	142 (10.1)	
Stage III-IV	7716 (24.2)	458 (32.6)	<.01
Recurrent cancer, n (%)	2161 (6.8)	82 (5.8)	.19
Comorbidities, n (%)			

Variable	No worse discharge (n=31,949)	Worse discharge (n=1406)	*P* value
CCI[d] ≥3	4508 (14.1)	227 (16.1)	.04
HFRS[e] ≥5	270 (0.8)	55 (3.9)	<.01
Cerebrovascular disease	853 (2.7)	87 (6.2)	<.01
Chronic pulmonary disease	1832 (5.7)	95 (6.8)	.12
Congestive heart failure	823 (2.6)	72 (5.1)	<.01
Dementia	1783 (5.6)	403 (28.7)	<.01
Diabetes	4122 (12.9)	232 (16.5)	<.01
Liver disease	987 (3.1)	66 (4.7)	<.01
Myocardial infarction	223 (0.7)	17 (1.2)	.04
Peptic ulcer disease	1597 (5)	85 (6)	.09
Peripheral vascular disease	195 (0.6)	16 (1.1)	.02
Renal disease	341 (1.1)	27 (1.9)	<.01
Medical history within 8 weeks, n (%)			
Chemotherapy	2573 (8.1)	88 (6.3)	.02
Radiation	75 (0.2)	6 (0.4)	.25
Surgery	4220 (13.2)	215 (15.3)	.03
BT[f] ≥38˚C, n (%)	4212 (13.2)	190 (13.5)	.75
sBP[g] ≥180 mm Hg, n (%)	1898 (5.9)	126 (9)	<.01
Albumin <3.5 g/dL, n (%)	6669 (20.9)	543 (38.6)	<.01
T-Bil[h] ≥2.0 mg/dL, n (%)	252 (0.8)	14 (1)	.48
Creatinine ≥2.0 mg/dL, n (%)	639 (2)	65 (4.6)	<.01
Platelet $<10^5$/μL, n (%)	641 (2)	59 (4.2)	<.01
Hemoglobin <11 g/dL, n (%)	5896 (18.5)	523 (37.2)	<.01
Number of beds, n (%)			<.01
<300	1298 (4.1)	70 (5)	
300 - 499	17,922 (56.1)	707 (50.3)	
≥500	12,729 (39.8)	629 (44.7)	

[a]Low income" means an estimated household income in the lowest tertile based on ZIP code.

[b]Barthel Index ≤ 60" means the Barthel Index ≤ 60 at admission.

[c]With epidural anesthesia" means the combination of general and epidural anesthesia.

[d]CCI: Charlson Comorbidity Index.

[e]HFRS: Hospital Frailty Risk Score.

[f]BT: body temperature.

[g]sBP: systolic blood pressure.

[h]T-Bil: total bilirubin.

Figure 3. Receiver operating characteristic curves of 6 machine-learning models in the training set. AUC: area under the receiver operating characteristic curve; CatBoost: category boosting; Logistic: logistic regression; NN: neural networks; RF: random forest; SVM: support vector machine; XGBoost: extreme gradient boosting.

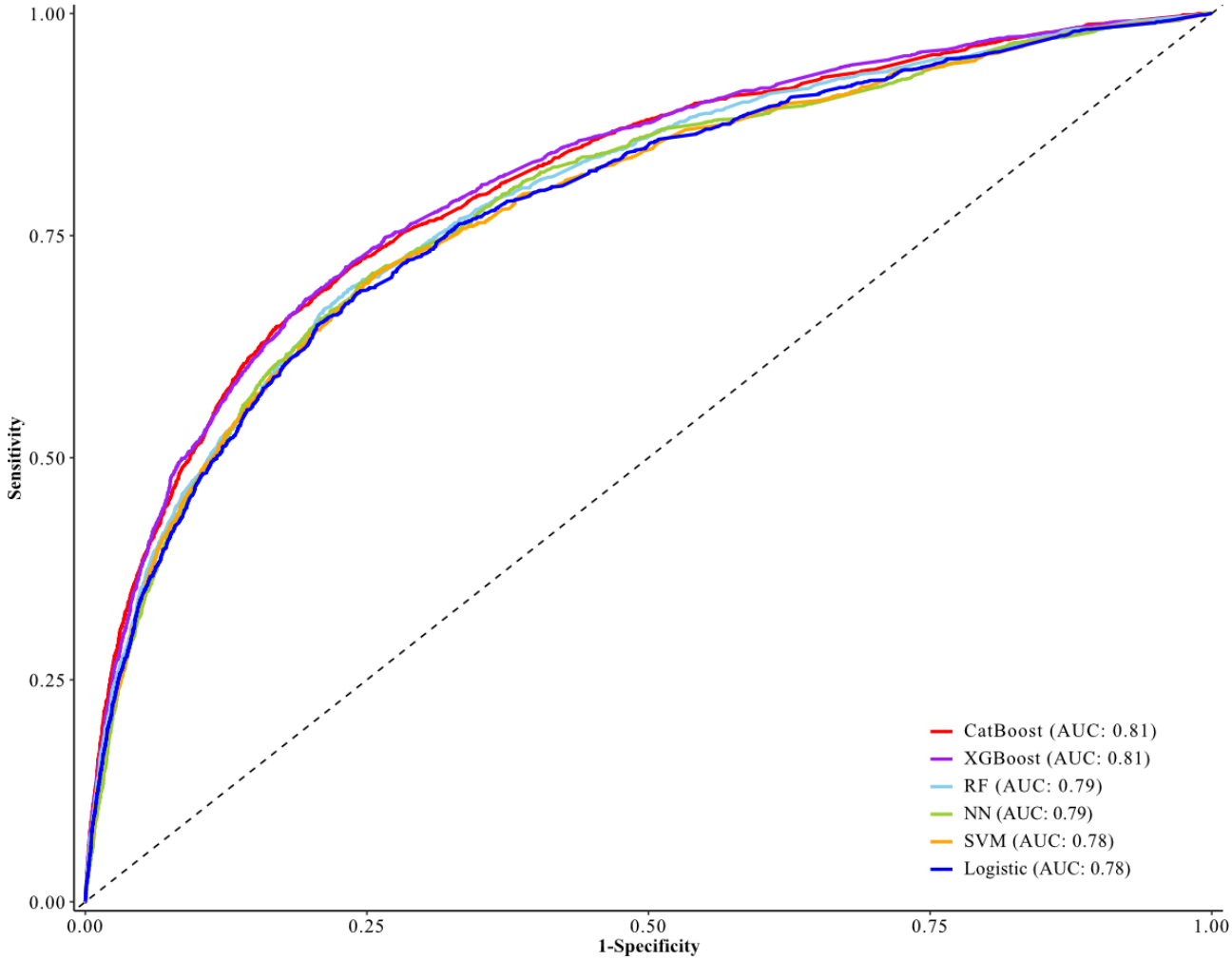


Figure 4. The top 15 features of predictor variables based on mean absolute Shapley additive explanations value. (**A**) The top 15 features selected by the category boosting model. (**B**) The top 15 features selected by the XGBoost: extreme gradient boosting model. (a) “Low income” means an estimated household income in the lowest tertile based on ZIP code. (b) “Barthel Index ≤60” means the Barthel Index ≤60 at admission. (c) “Chemotherapy” means that patients underwent chemotherapy within 8 weeks before surgery. (d) “Stage III-IV” means that patients had cancer staging III or IV using the TNM: Tumor Nodes Metastasis classification system.CatBoost: category boosting; SHAP: Shapley additive explanations; XGBoost: extreme gradient boosting.

Figure 5. Receiver operating characteristic curves of 6 machine learning models. (**A**) Receiver operating characteristic curves in the internal validation set. (**B**) Receiver operating characteristic curves in the external validation set. AUC: area under the receiver operating characteristic curve; CatBoost: category boosting; Logistic: logistic regression; NN: neural networks; RF: random forest; ROC: receiver operating characteristic; SVM: support vector machine; XGBoost: extreme gradient boosting.

Table . Adverse outcome between low-risk and high-risk group by category boosting model using the Youden index as cut-off value.

Variable	Internal validation set			External validation set		
	Low risk (n=10,788)	High risk (n=3506)	*P* value	Low risk (n=5459)	High risk (n=1252)	*P* value
Worse discharge, n (%)	221 (2)	386 (11)	<.01	116 (2.1)	113 (9)	<.01
Death, n (%)	25 (0.2)	33 (0.9)	<.01	16 (0.3)	18 (1.4)	<.01
Functional disability, n (%)[a]	196 (1.8)	353 (10.1)	<.01	100 (1.8)	95 (7.6)	<.01
LOS[b], median (IQR), days	9 (7-13)	13.0 (9-19)	<.01	10 (7-14)	13 (10-19)	<.01
Cost, median (IQR), US $						
Total	10,092 (7894-11,893)	11,048 (9191-13,106)	<.01	10,371 (8820-11,936)	11,069 (9401-13,499)	<.01
Medical consultation	77 (49-114)	97 (65-141)	<.01	75 (48-110)	97 (70-137)	<.01
Medication	68 (33, 139)	176 (87-362)	<.01	67 (34-139)	146 (76-333)	<.01
Medical procedure	26 (15-42)	40 (24-67)	<.01	27 (14-44)	28 (13-57)	<.01
Surgical procedure	6645 (5028-8253)	6620 (5513-8116)	.02	7203 (5825-8176)	6924 (5900-8212)	.46
Laboratory tests	449 (328-597)	516 (370-742)	<.01	443 (328-610)	526 (371-775)	<.01
Hospital stays	2404 (1873-3021)	3085 (2362-4073)	<.01	2484 (1940-3078)	3039 (2419-4178)	<.01
Others	0 (0-137)	100 (0-212)	<.01	0 (0-120)	0 (0-177)	<.01

[a]Functional disability" means a decrease in the Barthel Index by ≥5 points at discharge compared with admission.

[b]LOS: length of stay.

Discussion

Principal Findings

In this study, we developed and validated machine-learning models to predict postoperative functional disability and mortality in older patients with cancer. Our CatBoost model achieved good performance using routinely available preoperative factors from electronic health records, indicating the potential for clinical implementation. Although ethical training for hospital staff is essential to prevent unauthorized disclosure of prediction results, implementing this model within closed electronic health record systems could provide protection for patient privacy.

The previous model for lower-extremity surgery had an AUC of 0.72 in external validation [6], similar to our model; however, our model directly predicted functional disability using the Barthel Index rather than using nursing home discharge as a surrogate. The model performance remained consistent across sensitivity analyses for death and functional disability separately, and complete cases, indicating the robustness of our findings. Notably, the model demonstrated higher predictive performance in patients with stomach or colorectal cancer than the other cancers, making it especially valuable for surgical decision-making in these patients.

The CatBoost model identified patients at high risk who had longer LOS and higher health care costs; however, surgical procedure costs were comparable between patients at high and low risk. These findings suggested that the increased cost was based on the varying postoperative course. Our model can support decision-making for older patients with cancer and their families regarding cancer surgery by providing insights into potential postoperative QOL and care burden. Moreover, if patients at high risk choose to undergo cancer surgery, our model may enable health care providers to implement targeted interventions such as intensive postoperative rehabilitation. Early identification of patients at high risk, such as those aged ≥85 years with dementia, can help health care providers prepare support systems, including caregiver education, social work consultation for home health support, and coordination with multidisciplinary teams [15]. This proactive approach may help reduce caregiver burden and improve outcomes for both patients and their families.

Of the 6 machine-learning models, the CatBoost and XGBoost models, with AUC ≥0.80, had the same combination of 14 features in the top 15 influential factors. These factors include established risk factors for poor postoperative outcomes in older patients as identified in previous studies: dementia [7], older age [9] (≥85 y and 75 - 84 y), male sex [24], anemia [10] (hemoglobin <11 g/dL), low income [6], underweight [12] (BMI <18.5 kg/m²), diabetes [49], and cancer staging [25]. In addition,

several factors serve as proxies for known risk factors. For frailty [5], the factors include (1) open surgery, which generally results in a more pronounced postoperative functional disability compared with scopic surgeries; (2) nonhome admission, likely indicating that patients are too frail to live independently; and (3) Barthel Index ≤60 at admission, indicating severe dependence [22]. For malnutrition [13], the proxies are (1) albumin <3.5 g/dL, a marker of malnutrition, and (2) gastrointestinal cancer, which often involves a long time to restart food intake after surgery, increasing the risk of malnutrition compared with other cancer types. The consistency between the 2 models in identifying these factors further validates their importance in predicting postoperative outcomes. Although chemotherapy and creatinine ≥2.0 mg/dL were not common in both models, these factors, identified as influential factors in previous studies [33,50], might also be included as predictor variables in future models. Moreover, our models included several factors associated with social vulnerability, such as age ≥85 years, dementia, and low income. Without ensuring model transparency to health care providers, our model could unconsciously contribute to reduced surgical care access for vulnerable populations. Therefore, when implementing our model in clinical practice, health care providers should consider these model characteristics to ensure fair allocation of health care resources [51].

Limitations

Our study has some limitations. First, we only validated all models using Japanese data. While our CatBoost model showed moderate accuracy (AUC: 0.7 - 0.9) [52] in both internal and external validation, the AUC in the external validation was lower than that in the internal validation as observed in a previous study [6]. Further studies using data from other countries and ethnic groups are necessary to evaluate model robustness, including potential bias and overfitting [51], and confirm their applicability to different health care systems. However, our study used data from 70 hospitals across Japan, which may enhance generalizability within the country. In addition, considering the global trend of population aging, our models may prove valuable for other countries in the future, particularly when these countries reach levels of demographic aging similar to Japan's current situation.

Second, we did not have information on predictor variables such as marital status [6] because of the retrospective nature of the study. Despite this limitation, our models had good predictive performance in the validation sets. While our analysis showed no significant interactions between stage III–IV cancer and the top features, future studies incorporating additional variables may evaluate such interactions to enhance the predictive performance of models.

Finally, the long-term prognosis of patients classified as high-risk by our models remains unclear. Further research is required to determine the extent of functional recovery and mortality in these patients. At a minimum, postoperative functional disability in patients at high risk indicates an increased immediate post-discharge burden on family caregivers and health care resources.

Conclusions

Our CatBoost model achieved good performance for predicting postoperative functional disability and mortality in older patients with cancer. This model could support surgical decision-making for patients and families while guiding targeted interventions by health care providers. This model, which is based on routinely available preoperative factors, has the potential for implementation in clinical settings through electronic health records.

Acknowledgments

We thank Masato Koizumi of the Information Technology Department in the National Hospital Organization Headquarters for his advice on data handling. This research was partially funded by the Daiwa Securities Foundation, JSPS KAKENHI (grant number 23K19882), and Pfizer Health Research Foundation.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to being sensitive personal information. The analytic code is available from the corresponding author on request.

Authors' Contributions

YH, NI, TT, and SI contributed to conception and design and interpretation of data. YH assisted with drafting of the manuscript. YH, NI, TT, and SI handled critical review of the manuscript for important intellectual content. YH and NI contributed to statistical analysis. YH obtained funding. TT and SI assisted with supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table S1. ICD-10 codes selected as predictor variables. Table S2. Crude odds ratios of predictor variables for worse discharge. Table S3. Performance metrics of six machine learning models in training set. Table S4. Interaction between stage III–IV and top 5 features of predictor variables based on mean absolute SHAP value in training set. Table S5. AUCs of six machine learning

models for internal and external validation set in sensitivity and subgroup analyses. Figure S1. Precision-recall curve of six machine learning models in training set.
[DOCX File, 185 KB - aging_v8i1e65898_app1.docx]

References

1. Global cancer burden growing, amidst mounting need for services. World Health Organization. 2024. URL: https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services [accessed 2025-05-06]
2. Cancer statistics. National Cancer Center, Center for Cancer Control and Information Services. 2019. URL: https://www.ncc.go.jp/en/cis/index.html [accessed 2025-05-06]
3. World population prospects. United Nations Department of Economic and Social Affairs Population Division. 2019. URL: https://population.un.org/wpp/assets/Files/WPP2019_Highlights.pdf [accessed 2025-05-06]
4. Cabinet Office Japan. Annual report on the ageing society [summary]. 2021 Jul URL: https://www8.cao.go.jp/kourei/english/annualreport/2021/pdf/2021#:~:text=population%20aged%2075%20years%20and,65%2D74%20years%20in%20size.&text=By%202065%2C%20one%20in%202.6,75%20years%20old%20and%20over [accessed 2025-05-06]
5. Shaw JF, Mulpuru S, Kendzerska T, et al. Association between frailty and patient outcomes after cancer surgery: a population-based cohort study. Br J Anaesth 2022 Mar;128(3):457-464. [doi: 10.1016/j.bja.2021.11.035] [Medline: 35034792]
6. Schaefer MS, Hammer M, Platzbecker K, et al. What factors predict adverse discharge disposition in patients older than 60 years undergoing lower-extremity surgery? The adverse discharge in older patients after lower-extremity surgery (ADELES) risk score. Clin Orthop Relat Res 2021 Mar 1;479(3):546-547. [doi: 10.1097/CORR.0000000000001532] [Medline: 33196587]
7. Masutani R, Pawar A, Lee H, Weissman JS, Kim DH. Outcomes of common major surgical procedures in older adults with and without dementia. JAMA Netw Open 2020 Jul 1;3(7):e2010395. [doi: 10.1001/jamanetworkopen.2020.10395] [Medline: 32662842]
8. Morishima T, Kuwabara Y, Saito MK, et al. Patterns of staging, treatment, and mortality in gastric, colorectal, and lung cancer among older adults with and without preexisting dementia: a Japanese multicentre cohort study. BMC Cancer 2023 Jan 19;23(1):67. [doi: 10.1186/s12885-022-10411-y] [Medline: 36658524]
9. Konishi T, Sasabuchi Y, Matsui H, Tanabe M, Seto Y, Yasunaga H. Long-term risk of being bedridden in elderly patients who underwent oncologic surgery: a retrospective study using a Japanese claims database. Ann Surg Oncol 2023 Aug;30(8):4604-4612. [doi: 10.1245/s10434-023-13566-5] [Medline: 37149549]
10. Boening A, Boedeker RH, Scheibelhut C, Rietzschel J, Roth P, Schönburg M. Anemia before coronary artery bypass surgery as additional risk factor increases the perioperative risk. Ann Thorac Surg 2011 Sep;92(3):805-810. [doi: 10.1016/j.athoracsur.2011.02.076] [Medline: 21592459]
11. Schaap LA, Koster A, Visser M. Adiposity, muscle mass, and muscle strength in relation to functional decline in older persons. Epidemiol Rev 2013;35(35):51-65. [doi: 10.1093/epirev/mxs006] [Medline: 23221972]
12. Roh L, Braun J, Chiolero A, et al. Mortality risk associated with underweight: a census-linked cohort of 31,578 individuals with up to 32 years of follow-up. BMC Public Health 2014 Apr 16;14(14):371. [doi: 10.1186/1471-2458-14-371] [Medline: 24739374]
13. Kim SW, Han HS, Jung HW, et al. Multidimensional frailty score for the prediction of postoperative mortality risk. JAMA Surg 2014 Jul;149(7):633-640. [doi: 10.1001/jamasurg.2014.241] [Medline: 24804971]
14. Minami T, Fujita K, Hashimoto M, et al. External beam radiotherapy combination is a risk factor for bladder cancer in patients with prostate cancer treated with brachytherapy. World J Urol 2023 May;41(5):1317-1321. [doi: 10.1007/s00345-023-04380-5] [Medline: 37024557]
15. Huang LW, Smith AK, Wong ML. Who will care for the caregivers? Increased needs when caring for frail older adults with cancer. J Am Geriatr Soc 2019 May;67(5):873-876. [doi: 10.1111/jgs.15863] [Medline: 30924526]
16. Loyd C, Markland AD, Zhang Y, et al. Prevalence of hospital-associated disability in older adults: a meta-analysis. J Am Med Dir Assoc 2020 Apr;21(4):455-461. [doi: 10.1016/j.jamda.2019.09.015] [Medline: 31734122]
17. Ogawa M, Yoshida N, Nakai M, et al. Hospital-associated disability and hospitalization costs for acute heart failure stratified by body mass index- insight from the JROAD/JROAD-DPC database. Int J Cardiol 2022 Nov 15;367(367):38-44. [doi: 10.1016/j.ijcard.2022.08.044] [Medline: 36029847]
18. Kanazawa N, Tani T, Imai S, Horiguchi H, Fushimi K, Inoue N. Existing data sources for clinical epidemiology: database of the National Hospital Organization in Japan. Clin Epidemiol 2022;14(14):689-698. [doi: 10.2147/CLEP.S359072] [Medline: 35615723]
19. Hayashida K, Murakami G, Matsuda S, Fushimi K. History and profile of diagnosis procedure combination (DPC): development of a real data collection system for acute inpatient care in Japan. J Epidemiol 2021 Jan 5;31(1):1-11. [doi: 10.2188/jea.JE20200288] [Medline: 33012777]
20. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol 2017 Oct;27(10):476-482. [doi: 10.1016/j.je.2016.09.009] [Medline: 28142051]

21. Cancer statistics, annual report. National Cancer Center Japan. 2022. URL: https://www.ncc.go.jp/en/publication_report/2022/ncphs/ncphs14.html [accessed 2025-05-06]
22. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [doi: 10.1016/0895-4356(89)90065-6] [Medline: 2760661]
23. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3(1):32-35. [doi: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3] [Medline: 15405679]
24. Losurdo P, Mastronardi M, de Manzini N, Bortul M. Survival and long-term surgical outcomes after colorectal surgery: are there any gender-related differences? Updates Surg 2022 Aug;74(4):1337-1343. [doi: 10.1007/s13304-022-01323-4] [Medline: 35810269]
25. Bliton J, Parides M, Muscarella P, McAuliffe JC, Papalezova K, In H. Clinical stage of cancer affects perioperative mortality for gastrointestinal cancer surgeries. J Surg Res 2021 Apr;260(1-9):1-9. [doi: 10.1016/j.jss.2020.11.023] [Medline: 33310353]
26. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
27. Birim O, Kappetein AP, Bogers AJJC. Charlson comorbidity index as a predictor of long-term outcome after surgery for nonsmall cell lung cancer. Eur J Cardiothorac Surg 2005 Nov;28(5):759-762. [doi: 10.1016/j.ejcts.2005.06.046] [Medline: 16157485]
28. Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a hospital frailty risk score focusing on older people in acute care settings using electronic hospital records: an observational study. Lancet 2018 May 5;391(10132):1775-1782. [doi: 10.1016/S0140-6736(18)30668-8] [Medline: 29706364]
29. Cats A, Jansen EPM, van Grieken NCT, et al. Chemotherapy versus chemoradiotherapy after surgery and preoperative chemotherapy for resectable gastric cancer (CRITICS): an international, open-label, randomised phase 3 trial. Lancet Oncol 2018 May;19(5):616-628. [doi: 10.1016/S1470-2045(18)30132-3] [Medline: 29650363]
30. Erlandsson J, Holm T, Pettersson D, et al. Optimal fractionation of preoperative radiotherapy and timing to surgery for rectal cancer (Stockholm III): a multicentre, randomised, non-blinded, phase 3, non-inferiority trial. Lancet Oncol 2017 Mar;18(3):336-346. [doi: 10.1016/S1470-2045(17)30086-4] [Medline: 28190762]
31. Kovacs C, Jarvis SW, Prytherch DR, et al. Comparison of the National Early Warning Score in non-elective medical and surgical patients. Br J Surg 2016 Sep;103(10):1385-1393. [doi: 10.1002/bjs.10267] [Medline: 27487317]
32. Hartle A, McCormack T, Carlisle J, et al. The measurement of adult blood pressure and management of hypertension before elective surgery: Joint Guidelines from the Association of Anaesthetists of Great Britain and Ireland and the British Hypertension Society. Anaesthesia 2016 Mar;71(3):326-337. [doi: 10.1111/anae.13348] [Medline: 26776052]
33. Park JH, Lee HJ, Oh SY, et al. Prediction of postoperative mortality in patients with organ failure after gastric cancer surgery. World J Surg 2020 May;44(5):1569-1577. [doi: 10.1007/s00268-020-05382-9] [Medline: 31993720]
34. Brierley JD, Gospodarowicz MK, Wittekind C. TNM Classification of Malignant Tumours: John Wiley & Sons; 2017. [doi: 10.1002/9780471420194]
35. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005 Nov;43(11):1130-1139. [doi: 10.1097/01.mlr.0000182534.19832.83] [Medline: 16224307]
36. Liu C, Zhang K, Yang X, et al. Development and validation of an explainable machine learning model for predicting myocardial injury after noncardiac surgery in two centers in China: Retrospective Study. JMIR Aging 2024 Jul 26;7:e54872. [doi: 10.2196/54872] [Medline: 39087583]
37. Guan C, Ma F, Chang S, Zhang J. Interpretable machine learning models for predicting venous thromboembolism in the intensive care unit: an analysis based on data from 207 centers. Crit Care 2023 Oct 24;27(1):406. [doi: 10.1186/s13054-023-04683-4] [Medline: 37875995]
38. Mayer M. Package missranger: fast imputation of missing values. The Comprehensive R Archive Network. 2021 Mar 30. URL: https://cran.r-project.org/web/packages/missRanger/missRanger.pdf [accessed 2025-05-06]
39. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
40. Li Y, Salmasian H, Vilar S, Chase H, Friedman C, Wei Y. A method for controlling complex confounding effects in the detection of adverse drug reactions using electronic health records. J Am Med Inform Assoc 2014;21(2):308-314. [doi: 10.1136/amiajnl-2013-001718] [Medline: 23907285]
41. Chen Q, Mao R, Zhao J, et al. Nomograms incorporating preoperative RDW level for the prediction of postoperative complications and survival in colorectal liver metastases after resection. Ann Palliat Med 2021 Apr;10(4):4143-4158. [doi: 10.21037/apm-20-2418] [Medline: 33832316]
42. Konishi T, Goto T, Fujiogi M, et al. New machine learning scoring system for predicting postoperative mortality in gastroduodenal ulcer perforation: A study using A Japanese nationwide inpatient database. Surgery 2022 Apr;171(4):1036-1042. [doi: 10.1016/j.surg.2021.08.031] [Medline: 34538648]
43. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. arXiv. Preprint posted online on Jun 28, 2017. [doi: 10.48550/arXiv.1706.09516]

44. Tu JV. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J Clin Epidemiol 1996 Nov;49(11):1225-1231. [doi: 10.1016/s0895-4356(96)00002-9] [Medline: 8892489]
45. Oh T, Kim D, Lee S, et al. Machine learning-based diagnosis and risk factor analysis of cardiocerebrovascular disease based on KNHANES. Sci Rep 2022 Feb;12(1):35145205. [doi: 10.1038/s41598-022-06333-1]
46. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; Dec 4-9, 2017; Long Beach, CA. [doi: 10.48550/arXiv.1705.07874]
47. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [doi: 10.2307/2531595] [Medline: 3203132]
48. Bur AM, Brant JA, Newman JG, et al. Incidence and risk factors for prolonged hospitalization and readmission after transsphenoidal pituitary surgery. Otolaryngol Head Neck Surg 2016 Oct;155(4):688-694. [doi: 10.1177/0194599816652379] [Medline: 27221574]
49. Lopez LF, Reaven PD, Harman SM. Review: The relationship of hemoglobin A1c to postoperative surgical risk with an emphasis on joint replacement surgery. J Diabetes Complications 2017 Dec;31(12):1710-1718. [doi: 10.1016/j.jdiacomp.2017.08.016] [Medline: 29029935]
50. von Waldenfels G, Loibl S, Furlanetto J, et al. Outcome after neoadjuvant chemotherapy in elderly breast cancer patients - a pooled analysis of individual patient data from eight prospectively randomized controlled trials. Oncotarget 2018 Mar 16;9(20):15168-15179. [doi: 10.18632/oncotarget.24586] [Medline: 29632634]
51. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018 Dec 18;169(12):866-872. [doi: 10.7326/M18-1990] [Medline: 30508424]
52. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intensive Care Med 2003 Jul;29(7):1043-1051. [doi: 10.1007/s00134-003-1761-8] [Medline: 12734652]

Abbreviations

AUC: area under the receiver operating characteristic curve
CatBoost: category boosting
CCI: Charlson Comorbidity Index
HFRS: Hospital Frailty Risk Score
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*
Lasso: least absolute shrinkage and selection operator
LOS: length of stay
NHO: National Hospital Organization
OR: odds ratio
PRAUC: area under the precision-recall curve
QOL: quality of life
ROC: receiver operating characteristic
SHAP: Shapley additive explanations
TNM: Tumor Node Metastasis
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
XGBoost: extreme gradient boosting

Edited by M Mardini; submitted 29.08.24; peer-reviewed by H Li, KE Cain, YG Luo; revised version received 18.03.25; accepted 18.03.25; published 14.05.25.

Please cite as:
Hashimoto Y, Inoue N, Tani T, Imai S
Machine Learning for Predicting Postoperative Functional Disability and Mortality Among Older Patients With Cancer: Retrospective Cohort Study
JMIR Aging 2025;8:e65898
URL: https://aging.jmir.org/2025/1/e65898
doi:10.2196/65898

Original Paper

Unsupervised Deep Learning of Electronic Health Records to Characterize Heterogeneity Across Alzheimer Disease and Related Dementias: Cross-Sectional Study

Matthew West[1*], MPhys, SM; You Cheng[1*], PhD; Yingnan He[1*], MS; Yu Leng[1], SM; Colin Magdamo[1], BS; Bradley T Hyman[1], MD, PhD; John R Dickson[1], MD, PhD; Alberto Serrano-Pozo[1], MD, PhD; Deborah Blacker[1], MD, PhD; Sudeshna Das[1], PhD

Massachusetts General Hospital, Cambridge, MA, United States

[*]these authors contributed equally

Corresponding Author:
Sudeshna Das, PhD
Massachusetts General Hospital
65 Landsdowne Street
Cambridge, MA, 02139
United States
Phone: 1 617 768 8254
Email: sdas5@mgh.harvard.edu

Abstract

Background: Alzheimer disease and related dementias (ADRD) exhibit prominent heterogeneity. Identifying clinically meaningful ADRD subtypes is essential for tailoring treatments to specific patient phenotypes.

Objective: We aimed to use unsupervised learning techniques on electronic health records (EHRs) from memory clinic patients to identify ADRD subtypes.

Methods: We used pretrained embeddings of non-ADRD diagnosis codes (*International Classification of Diseases, Ninth Revision*) and large language model (LLM)–derived embeddings of clinical notes from patient EHRs. Hierarchical clustering of these embeddings was used to identify ADRD subtypes. Clusters were characterized regarding their demographic and clinical features.

Results: We analyzed a cohort of 3454 patients with ADRD from a memory clinic at Massachusetts General Hospital, each with a specialist diagnosis. Clustering pretrained embeddings of the non-ADRD diagnosis codes in patient EHRs revealed the following 3 patient subtypes: one with skin conditions, another with psychiatric disorders and an earlier age of onset, and a third with diabetes complications. Similarly, using LLM-derived embeddings of clinical notes, we identified 3 subtypes of patients as follows: one with psychiatric manifestations and higher prevalence of female participants (prevalence ratio: 1.59), another with cardiovascular and motor problems and higher prevalence of male participants (prevalence ratio: 1.75), and a third one with geriatric health disorders. Notably, we observed significant overlap between clusters from both data modalities (χ^2_4=89.4; P<.001).

Conclusions: By integrating *International Classification of Diseases, Ninth Revision* codes and LLM-derived embeddings, our analysis delineated 2 distinct ADRD subtypes with sex-specific comorbid and clinical presentations, offering insights for potential precision medicine approaches.

(JMIR Aging 2025;8:e65178) doi:10.2196/65178

KEYWORDS

Alzheimer disease and related dementias; electronic health records; large language models; clustering; unsupervised learning

Introduction

Background

Alzheimer disease (AD) is a neurodegenerative condition which affects more than 55 million people globally [1], and it is the seventh leading cause of death in the United States [2]. Despite its substantial public health burden, AD remains poorly understood, with limited treatment options available. AD and related dementias (ADRD) is an umbrella term that refers to multiple dementing illnesses, including AD, frontotemporal dementia (FTD), Lewy body dementia (LBD), and vascular dementia. AD is the most prevalent, accounting for around 60% to 80% of all dementia [2]. While these diseases have distinct clinical and neuropathological criteria, there is substantial overlap in both clinical presentation and autopsy findings at the individual patient level. For example, AD is clinically characterized by an amnestic-predominant dementia and neuropathologically defined by the build-up of amyloid beta (Aβ) plaques and neurofibrillary tangles formed by hyperphosphorylated tau protein [3]; however, these lesions are frequently accompanied by cerebrovascular disease (CVD) [4] or Lewy body pathology [5], which can influence clinical presentation. Likewise, LBD is defined by Lewy bodies but is also associated with plaques and tangles [6], which may accelerate the rate of cognitive decline [7]. This clinical and neuropathological heterogeneity limits our ability to target disease-modifying drugs to each specific neuropathological lesion. The so-called *amyloid hypothesis* has prevailed as the leading explanation of AD disease etiology, where it is held that Aβ toxicity leads to tau hyperphosphorylation, synaptic dysfunction, and neurodegeneration [8]. However, treatments targeting this hypothesis only show limited efficacy, which may stem in part from the clinical and neuropathological comorbidities [9], highlighting the need for a tailored approach to identify potential subtypes of disease and develop more effective targeted treatments.

Previous approaches to AD subtyping have focused on RNA expression, as well as brain imaging and cognitive assessments. Neff et al [10] identified 5 molecular subtypes of AD using RNA-sequencing signatures, characterized by different dysregulated pathways related to tau-mediated neurodegeneration, Aβ neuroinflammation, synaptic signaling, immune activity, mitochondria organization, and myelination. The Subtype and Stage Inference algorithm, applied to magnetic resonance imaging and positron emission tomography imaging data, identified distinct AD trajectories based on the rate and sequence of brain atrophy [11] and tau deposition [12]. Cognitive subtypes have also been identified based on memory, visuospatial and linguistic capabilities, and executive function [13-15]. These studies were limited to research cohorts with specific selection criteria, and it is unclear whether these subtypes can be extended to larger samples.

In contrast, real-world data, such as, electronic health records (EHRs), provide readily accessible large observational datasets and have been used for clustering AD or ADRD subtypes [16]. Unsupervised learning approaches on EHR datasets have revealed latent structure in conditions, such as autism [17,18] and Parkinson disease [19]. For AD subtyping, EHR-based approaches have used the *International Classification of Diseases* (*ICD*) or similar diagnostic codes, showing varying success depending on the methodology and population. Xu et al [20] used hierarchical clustering on EHR data from patients with AD, identifying subtypes related to CVD, mental illness, age of onset, and sensory problems. Alexander et al [21] found the following 5 patient subtypes: mental health, nontypical AD, typical AD, CVD, and men with cancer. They later identified a consistent subtype with early-onset AD, predominantly female participants, with a faster rate of progression using various machine learning methods [22]. Landi et al [23] used unsupervised deep learning to encode EHRs with temporal information, identifying early-onset AD, later-onset AD with mild comorbidities, and typical-onset AD with moderate symptoms. He et al [24] applied spectral clustering to EHRs of patients with AD, discerning 4 subtypes with significant demographic, mortality, and medication use differences. Tang et al [25] analyzed comorbidity patterns in EHRs of patients with AD, revealing sex-dependent variations. In another study, Tang et al [26] used EHRs with knowledge networks to predict AD onset and identify sex-specific genetic markers. These studies collectively highlight the varied methodologies and results in EHR-based research, emphasizing the complexity and potential of these approaches for a deeper understanding of AD.

However, none of these prior studies leveraged the richer representation of EHR data by embedding full sequences of clinical text. Transformers have emerged as state-of-the-art architecture for language modeling and are broadly characterized by the concept of attention [27]. Attention, named for its similarity to cognitive attention, enables the sharing of contextual information among word representations without directly encoding their sequence. The transformer used in this work is a version of the Bidirectional Encoder Representations from Transformers (BERT) architecture [28]. This architecture consists of an encoder which can be fine-tuned on downstream applications and domains. Specifically, we use Clinical BERT [29], which is pretrained on a large corpus of clinical notes from the critical care database Medical Information Mart for Intensive Care (MIMIC) [30].

Objectives

In this work, we used both pretrained embeddings of *ICD-9* code diagnostic data and transformer-derived embeddings of clinical notes. This dual approach addresses the limitations of previous studies by incorporating structured *ICD* codes, which allow us to study subtypes of patients with similar *ICD* codes (non-ADRD diagnosis in charts), and unstructured clinical notes, which capture detailed clinical history and manifestations provided by specialists. By combining these 2 modalities, we aimed to enhance the clustering of patient ADRD subtypes.

Methods

Cohort Selection Process

Patients were selected from the Massachusetts General Hospital (MGH) EHR database. The selection criteria included patients who had at least 2 MGH memory clinic visits (either an in-person office visit or a video telemedicine visit) from August

2015 to June 2022, were aged >50 years at their first visit, and had progress notes of substantial length (≥512 characters). These criteria were chosen due to the richness of the notes for the clustering analysis and the high quality of the ADRD diagnosis from specialists. From the identified patient cohort, 2 datasets were extracted as follows: one containing structured diagnostic *ICD* code data from the patients' entire medical history and another consisting of unstructured clinical notes authored by memory clinic specialists, limited to the most recent visit. We chose only the most recent visit note because it typically consolidates the patient's prior history, thereby reducing redundancy and providing a focused, up-to-date clinical snapshot. In addition, the dataset was filtered to exclude patients who did not have ADRD diagnoses (Multimedia Appendix 1), as well as those who lacked non-ADRD *ICD* codes (ie, patients who only had ADRD *ICD* codes were excluded).

Ethical Considerations

This study was approved by the Mass General Brigham Institutional Review Board (protocol 2015P001915), which granted a waiver of informed consent for secondary analysis of electronic health data. No participant compensation was provided. Electronic health data was queried from Epic and securely stored on servers within the Mass General Brigham firewall. Access was restricted to authorized study personnel, in full compliance with institutional privacy and data security policies.

Embedding Methodology

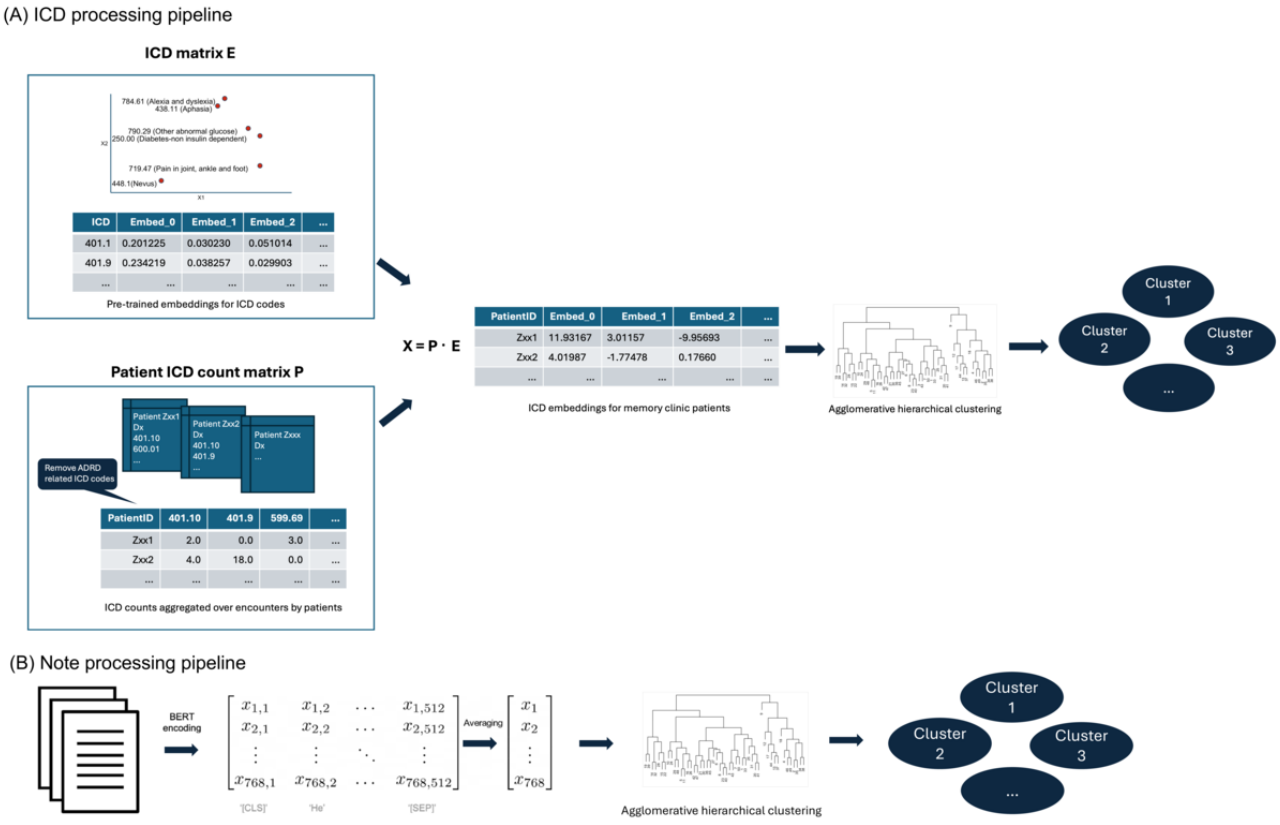
ICD Codes

Before clustering, it was necessary to derive a patient-level representation that encoded information relevant to phenotype in a single vector. While some prior work has relied on one-hot encoding (where categorical data are converted into binary vectors) of clinical data to represent patient phenotype, we leverage existing pretrained embeddings that capture relevant biomedical semantics in their latent representations of clinical concepts. In particular, we use a set of 300-dimensional embeddings for *ICD-9* codes, derived from prior work by Choi et al [31].

For a count-based encoded representation of *m ICD-9* codes across our cohort, $P \in R^{3454 \times m}$, and an embedding matrix, $E \in R^{m \times 300}$, our design matrix for clustering, $X \in R^{3454 \times 300}$, is given by the following matrix multiplication: $X = P \cdot E$.

This matrix multiplication sums the non-ADRD *ICD* embeddings across a patient record, and the resultant embedding is directly affected by the number of times each code appears in a patient's history. ADRD codes were dropped from the matrix *P* to not confound clustering based on structured ADRD phenotype. A schematic depicting the *ICD* representation pipeline is provided in Figure 1A.

Figure 1. Visualization of the clustering pipeline for (A) International Classification of Diseases (ICD) codes and (B) notes. For each subfigure, the workflow goes from left to right. BERT: Bidirectional Encoder Representations from Transformers.



Clinical Notes

Clinical notes were encoded using Clinical BERT before clustering. To derive patient-level representations of clinic notes, several preprocessing steps were undertaken. First, unwanted delimiter characters were stripped from patient notes, and notes were chunked into contiguous sections of up to 1024 characters. This resulted in a distribution of token numbers of ~200 to 300 per note following BERT's WordPiece encoding of input sequences. After passing through the transformer encoding, we took the final layer representation averaged over the 12 attention heads, such that each note was represented by a matrix of dimension (768, n), corresponding to each of the n input tokens having a 768-dimensional contextual vector representing it. Following this encoding, the representation was averaged over the token dimension to arrive at a single 768-dimensional vector for the whole note. This was explored using both simple averaging (arithmetic mean) over the token dimension, as well as attention-weighted averaging based on row-wise entropy of the final layer attention matrix. Attention-weighted averaging was used in patient representations due to the resultant lower inertia and increased silhouette score on average. A schematic depicting the note representation pipeline is provided in Figure 1B.

Attention-Weighted Averaging

For a given input sequence of length n, the final layer representation in a transformer model has an associated self-attention matrix, $A \in R^{n \times n}$. For BERT-based models, n is ≤512 due to the constraint on the length of the context window. To provide weights for averaging the embedding, $E \in R^{768 \times n}$, over the token dimension, we compute the row-wise differential entropy of the attention matrix, given in equation (1). Differential entropy is the continuous analog of Shannon entropy, which is usually only defined for discrete random variables [32]. In particular, the method described in Ebrahimi et al [33] is used to approximate the differential entropy, implemented in the Python (Python Software Foundation) library *SciPy version 1.7.3* [34], as the closed-form expression for the attention distribution for a given row $f(x)$ is not known analytically from the values of attention sampled. The differential entropy for a row i is given by

and the corresponding vector $h \in R^{n \times 1}$ corresponds to the entropy across every row. From the row-wise entropy, this vector is softmaxed to obtain the corresponding weights, $w \in R^{n \times 1}$, as follows:

The resultant embedding for a note sequence, $N \in R^{768 \times 1}$, is then given by the matrix multiplication as follows:

and the final patient-level representation is the simple average over all note fragments for a given patient, for their most recent encounter. A visualization of attention matrices with varying row-wise entropy and thus varying weighting per token is provided in Figures S1 and S2 in Multimedia Appendix 2.

Hierarchical Clustering

Clustering analysis was performed on *ICD-9* embeddings and clinical text representations to identify ADRD subtypes. We selected hierarchical agglomerative clustering with Ward linkage due to its ability to capture the hierarchical structure of clinical data, as seen in *ICD-9* codes (eg, metabolic disorders branching into type 1 and type 2 diabetes) and clinical notes (eg, cognitive impairment branching into memory loss and aphasia). Cluster quality was evaluated using elbow plots and silhouette scores, with implementation via *Scikit-learn v1.0.1* [35].

Optimal Transport

To address provider-specific effects in embeddings of clinical notes, we first applied Uniform Manifold Approximation and Projection (UMAP) to reduce the dimensionality of the data and then applied the earth mover distance transport approach using the Python *Optimal Transport* package [36]. *Optimal Transport* provides a mathematical framework to minimize the cost of transforming one distribution into another, which can address the problem of domain adaptation [37]. Domain adaptation involves adjusting data from different sources to make their data distributions more comparable, ensuring that models trained on these data perform well across various settings. In this context, the earth mover distance method was used to align the embeddings from various providers to a standard reference. This alignment ensured that the subsequent clustering analysis was less skewed by provider-related differences, allowing a more accurate interpretation of the underlying phenotypic variation.

Enrichment Analysis: ICD Clusters

ICD clusters were phenotypically characterized by testing for enrichment of *ICD-9* codes within each cluster. For each cluster, a 2×2 contingency table was generated for each *ICD-9* diagnosis code, comparing counts of patients with that code within cluster to counts of patients with that same code in other clusters. A chi-square test for enrichment was performed; the prevalence ratio (PR), calculated as the prevalence of each code in one cluster divided by its prevalence in the rest, was calculated to measure the strength of the association. A Bonferroni correction was applied to the resultant P values to correct for multiple comparisons, and *ICD* codes in each cluster were ranked by corrected P value to characterize the most significant enrichments. All P values in this investigation were two-sided, with a postcorrected α of 0.05 to determine significance. The top-10 significant diagnoses with the highest PR were extracted from each cluster for interpretation. If a cluster lacked significant diagnoses, the top diagnoses with the highest PR among the nonsignificant ones were selected. The enrichment analyses were conducted in Python version 3.8.15.

Topic Modeling: Note Clusters

We used BERTopic [38], using the Python package *BERTopic* v0.16.0, to identify representative topics and key terms within each note cluster. Embeddings obtained through optimal transport were directly used for clustering and topic assignment, bypassing the need for additional embedding and dimensionality

reduction steps. Before conducting cluster-based term frequency–inverse document frequency (TF-IDF) for topic assignment, the clinical text was preprocessed using a vectorizer to remove stop words and exclude common terms that appeared too frequently across most notes. Furthermore, to fine-tune and enhance the word representation of topics, we applied the KeyBERTInspired model, which extracts keywords by leveraging embeddings and cosine similarity to find the words with the closest semantic relationship to the note texts, thereby making them more representative of the topics. Following the extraction of representative terms within each cluster, we used GPT-4 (OpenAI) [39], a state-of-the-art large language model (LLM), to enhance interpretability. GPT-4 summarized the representative words provided by BERTopic into coherent themes with greater clinical significance, such as specific medical conditions, treatments, and medications. The PR, calculated as the prevalence of each word in one cluster divided by its prevalence in the rest, was calculated to measure the strength of the semantic relationship. The BERTopic modeling and analyses were conducted in Python version 3.9.6.

ADRD Diagnosis Categorization

The categorization of ADRD diagnoses was conducted using an extensive list of diagnostic names based on disease etiology. This list was meticulously reviewed for each unique diagnosis name recorded for the MGH memory clinic patients in the EHR system. The ADRD diagnoses categories included AD; dementia unspecified; FTD; LBD; vascular cognitive impairment (VCI); and others, such as posterior cortical atrophy (PCA), progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia. An expert behavioral neurologist (JRD) provided critical input during this process, helping to develop a comprehensive mapping list that correlates specific diagnosis names with their corresponding ADRD categories. The application of this mapping to the data was performed using R version 4.3.2 (R Foundation for Statistical Computing). The full list of diagnosis names corresponding to ADRD diagnosis categories is provided in Multimedia Appendix 1.

Cluster Characterizations

To assess associations between clusters and sex, as well as ADRD diagnoses, we used the chi-square test. For each cluster, a 2×2 contingency table was generated for each variable, comparing the counts of patients with the characteristic within the cluster to those in other clusters. The PR, defined as the prevalence of a characteristic in one cluster divided by its prevalence in the remaining clusters, was calculated to measure the strength of the association: 1 indicates no difference in prevalence between the 2 groups, >1 indicates higher prevalence in the first group, and <1 indicates lower prevalence in the first group. In addition, to examine variations in the age of onset across clusters, we initially conducted a Kruskal-Wallis Rank Sum Test using the *stats* package from R. η^2 (calculated by subtracting the number of groups from the Kruskal-Wallis H statistic plus one, and then dividing this result by the total number of observations minus the number of groups) based on the H statistic was reported as the effect size: values closer to 0 indicate a smaller effect and values closer to 1 indicate a larger effect. Following significant findings, further post hoc analyses using the Dunn test were performed to delineate differences between groups. The *P* values were adjusted for multiple comparisons using the Benjamini-Hochberg method to control the false discovery rate (FDR). Age-of-onset data were rigorously annotated by human experts reviewing clinical notes; where notes did not specify an exact age of onset, the age at the first clinical visit within the memory clinic was used as an approximation. Finally, we conducted a chi-square test between *ICD* clusters and note clusters to test whether patient cluster assignment was consistent across note and *ICD* representations. If the contingency table was larger than 2×2, Cramér V (calculated as the square root of the chi-square statistic divided by the product of the sample size and the minimum dimension minus one) was reported as the effect size: 0 indicates no association and 1 indicates a strong association. Standardized residuals (standardized differences between the observed count and the expected count) were reported for each cell: values close to 0 indicate the observed count is close to the expected count, positive values indicate the observed count is higher than expected, and negative values indicate the observed count is lower than expected. All statistical analyses were conducted in R version 4.3.2.

Results

Study Population

Our final study population consisted of 3454 patients from the MGH tertiary care memory clinic with clinical notes and *ICD* codes in the EHR system. The average age of onset for patients was 72.1 (SD 9.5) years, with 1678 (48.58%) being female. The majority were White (n=3059, 88.56%), followed by Asian (n=90, 2.61%), Black or African American (n=77, 2.23%), American Indian or Alaska Native (n=4, 0.12%), and Native Hawaiian or Other Pacific Islander (n=1, 0.03%). In addition, 103 (2.98%) identified as belonging to other races, and race data were not available for 120 (3.47%) patients. Regarding ethnicity, 3020 (87.43%) identified as non-Hispanic, 106 (3.07%) as Hispanic, and ethnicity data were not available for 328 (9.5%) patients. AD was the most prevalent diagnosis, affecting 1317 (38.13%) patients, followed by dementia unspecified, which accounted for 1101 (31.88%) patients. Each encounter recorded only one diagnosis name, and only the most recent encounter was used. The patient selection details are illustrated in Figure 2. The demographic and ADRD diagnosis breakdowns are provided in Table 1 and ADRD categorization details are provided in Multimedia Appendix 1.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram illustrating the selection of patients from the Massachusetts General Hospital (MGH) electronic health record (EHR) system. ADRD: Alzheimer disease and related dementias; Dx: diagnosis; ICD: International Classification of Diseases.

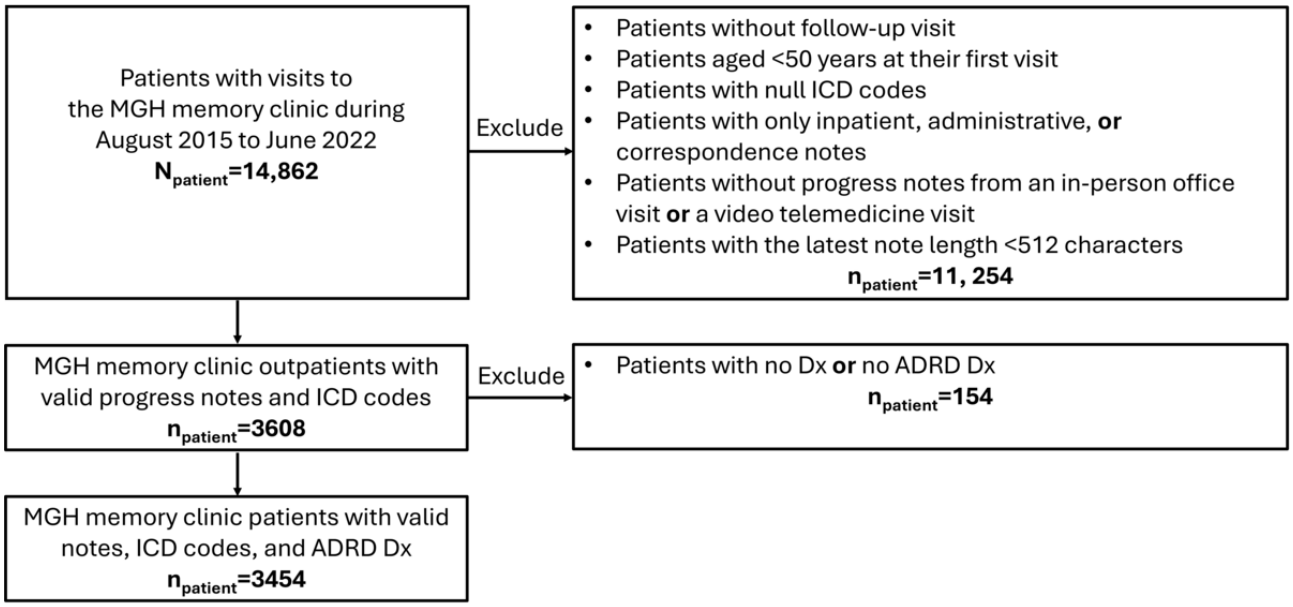


Table 1. Summary statistics of the final study population.

Characteristics	Total (N=3454)	AD[a] (n=1317)	Dementia unspecified (n=1101)	FTD[b] (n=190)	LBD[c] (n=261)	VCI[d] (n=96)	Other[e] (n=489)
Age of onset[f] (y), mean (SD)	72.1 (9.5)	74.6 (7.8)	73.3 (9.7)	63.8 (8.2)	71.4 (7.7)	76 (7.7)	65.1 (9.4)
Sex, n (%)							
Female	1678 (48.58)	717 (54.44)	538 (48.86)	78 (41.1)	64 (24.5)	41 (43)	240 (49.1)
Male	1776 (51.42)	600 (45.56)	563 (51.14)	112 (58.9)	197 (75.5)	55 (57)	249 (50.9)
Race, n (%)							
White	3059 (88.56)	1149 (87.24)	988 (89.74)	168 (88.4)	227 (87.0)	80 (83)	447 (91.4)
Black or African American	77 (2.23)	27 (2.05)	22 (2.00)	6 (3.2)	4 (1.5)	9 (9)	9 (1.8)
Asian	90 (2.61)	34 (2.58)	31 (2.82)	6 (3.2)	14 (5.4)	1 (1)	4 (0.8)
American Indian or Alaska Native	4 (0.12)	2 (0.15)	1 (0.09)	0 (0.0)	0 (0.0)	0 (0)	1 (0.2)
Native Hawaiian or Other Pacific Islander	1 (0.03)	1 (0.08)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other	103 (2.98)	52 (3.95)	32 (2.91)	2 (1.1)	2 (0.8)	3 (3)	12 (2.5)
Unavailable	120 (3.47)	52 (3.95)	27 (2.45)	8 (4.2)	14 (5.4)	3 (3)	16 (3.3)
Ethnicity, n (%)							
Not Hispanic or Latino	3020 (87.43)	1133 (86.03)	987 (89.65)	161 (84.7)	228 (87.4)	83 (87)	428 (87.5)
Hispanic or Latino	106 (3.07)	53 (4.02)	37 (3.36)	2 (1.1)	3 (1.1)	3 (3)	8 (1.6)
Unavailable	328 (9.50)	131 (9.95)	77 (6.99)	27 (14.2)	30 (11.5)	10 (10)	53 (10.8)

[a]AD: Alzheimer disease.

[b]FTD: frontotemporal dementia.

[c]LBD: Lewy body dementia.

[d]VCI: vascular cognitive impairment.

[e]Includes posterior cortical atrophy, progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia.

[f]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

ICD Clustering

To investigate the clinical heterogeneity within ADRD, we clustered the embeddings of non-ADRD *ICD* codes assigned to this tertiary care sample from the MGH ADRD cohort. We hypothesized that this approach would reveal distinct clinical subtypes based on clinical comorbidities, which were associated with demographics (ie, age of onset and sex). The hierarchical agglomerative clustering method revealed 3 distinct clusters in the embeddings of non-ADRD *ICD* codes, as determined by the silhouette score. Figure 3A depicts a heatmap of enriched *ICD* codes across each cluster, and a 2D UMAP projection of these *ICD* code embeddings, colored by cluster, is displayed in Figure 4A. The distribution of patients across the clusters was as follows: cluster 1 included 1501 (43.46%) patients, cluster 2 included 1597 (46.24%) patients, and cluster 3 included 356 (10.31%) patients. Detailed summary statistics for these clusters are presented in Table 2.

Figure 3. Heatmap of enrichment in International Classification of Diseases (ICD) clusters and topics in note clusters. (A) This heatmap displays the enrichment of ICD-9 codes across ICD embedding clusters. Cluster 1 is primarily dominated by skin-related and certain cardiovascular conditions. Cluster 2 is marked by its exclusive and high prevalence ratios (PR) in psychiatric and behavioral conditions. Cluster 3 shows a diverse set of conditions with a significant prevalence of respiratory, pain-related, and complicated diabetic mellitus. (B) This heatmap displays representative words for each note cluster identified through topic modeling. Cluster 1 is primarily dominated by psychiatric manifestations and medications. Cluster 2 highlights cardiovascular, motor, and sensory issues. Cluster 3 covers a variety of symptoms and conditions, including autoimmune issues, behavioral and movement disorders, sleep disturbances, etc. In both (A) and (B), the color intensity of each code or word-cluster pairing reflects the PR (prevalence of code or word in observed group divided by prevalence in other groups) associated with that code or word. Words colored as exclusive were only present in one cluster.

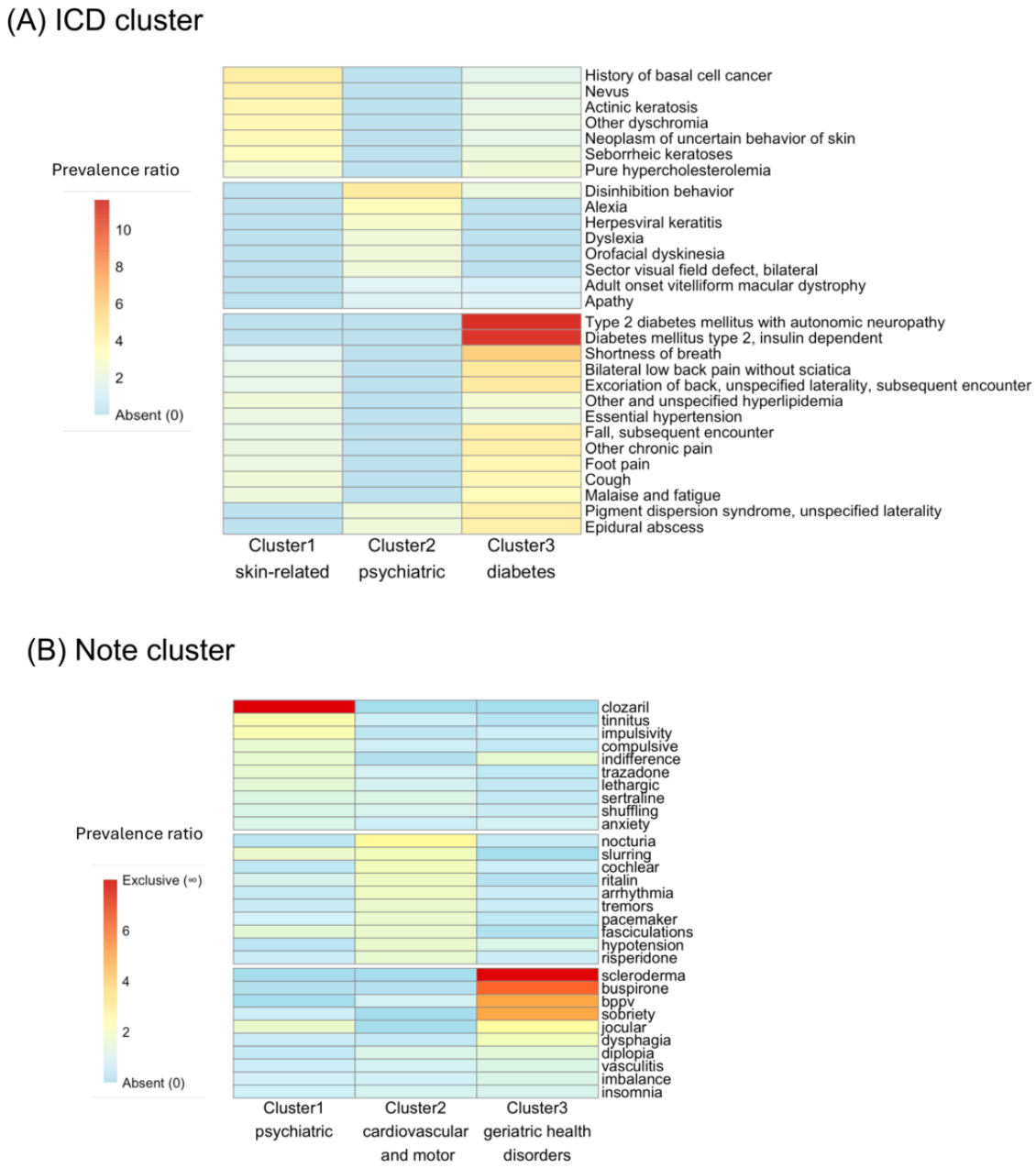
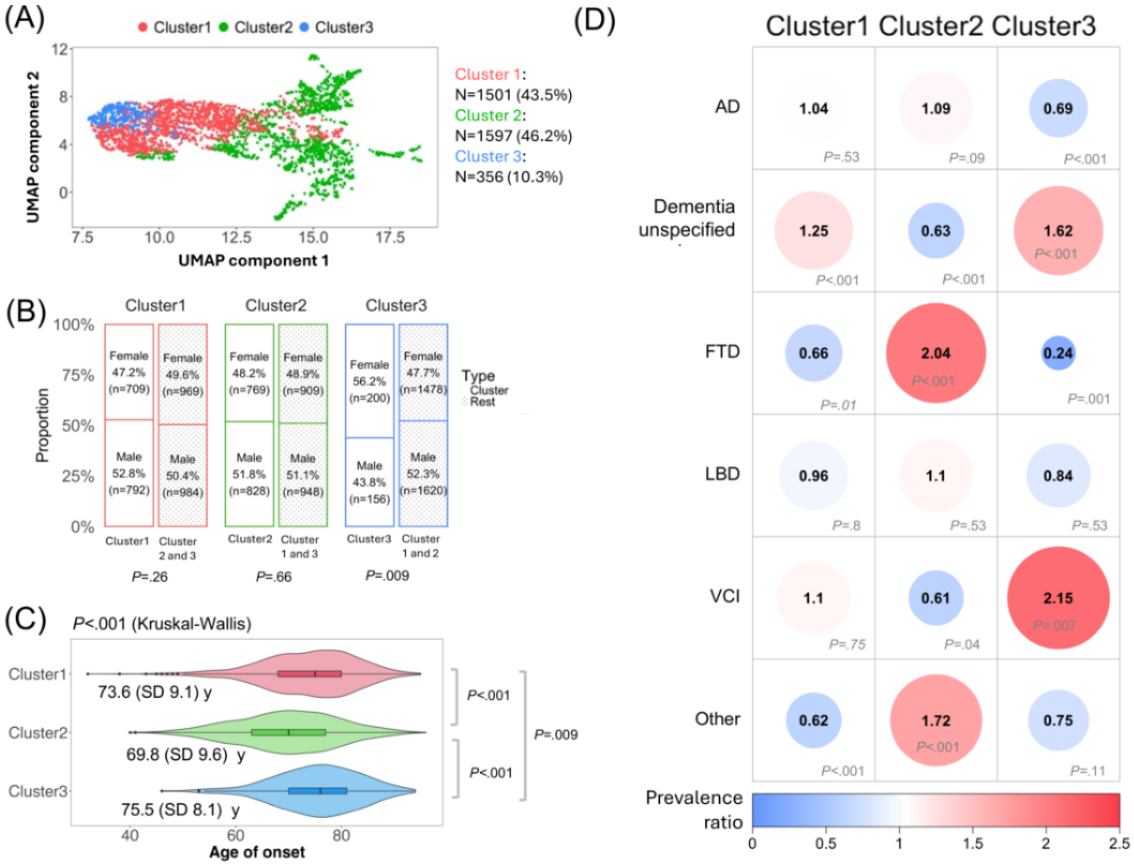


Figure 4. Clustering of International Classification of Diseases (ICD) embeddings and their demographic and diagnostic associations. (A) Uniform Manifold Approximation and Projection (UMAP) visualization of ICD embeddings were characterized by 3 clusters: cluster 1 includes 1501 (43.5%) patients, cluster 2 comprises 1597 (46.2%) patients, and cluster 3 contains 356 (10.3%) patients. (B) Bar plot showing prevalence for sex by cluster, significance based on a chi-square test. Notably, cluster 3 has a significantly higher proportion of female participants compared to male participants (P_{FDR}=.009). (C) Violin plot illustrating the distribution of age of onset across clusters. Each violin plot shows the kernel density estimate of the data, with the center line representing the median age of onset. Box plot elements are overlaid, where the box limits indicate the upper and lower quartiles, and the whiskers extend to 1.5 times the IQR. Individual points are hidden for clarity. Significant differences are observed, with cluster 2 showing the earliest average age of onset at 69.8 (SD 9.6) years, and cluster 3 the latest at 75.5 (SD 8.1) years (P<.001). (D) Heatmap showing prevalence ratio for Alzheimer disease and related dementias (ADRD) diagnoses across clusters, significance derived from a chi-squared test. Clinical diagnoses include Alzheimer disease (AD); dementia unspecified; frontotemporal dementia (FTD); Lewy body dementia (LBD); vascular cognitive impairment (VCI); and others such as posterior cortical atrophy (PCA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and primary progressive aphasia (PPA). Significant distribution variations are evident across clusters. The circle size, color, and number indicate the magnitude of the prevalence ratio. Blue indicates underrepresentation while red indicates overrepresentation. *P* values were corrected with false discovery rate adjustments.



We examined differences in non-ADRD *ICD* code frequency across patient *ICD* embedding clusters (Figure 3A). In cluster 1, diagnoses, such as seborrheic keratoses (χ^2_1=243.15; P_{FDR}<.001; PR=3.32), actinic keratosis (χ^2_1=236.53; P_{FDR}<.001; PR=3.91), pure hypercholesterolemia (χ^2_1=199.13; P_{FDR}<.001; PR=2.71), history of basal cell cancer (χ^2_1=196.22; P_{FDR}<.001; PR=4.45), and nevus (χ^2_1=184.53; P_{FDR}<.001; PR=4.23), show notably high PRs. These diagnoses largely fall into skin-related disorders (such as various types of skin cancer and keratoses). Cluster 2 appears to be unique, with top PRs noted only for clinical signs, such as disinhibition behavior (χ^2_1=1.14; P_{FDR}=.99; PR=4.65), alexia (χ^2_1=0.426; P_{FDR}=.99; PR=3.49), and orofacial dyskinesia (χ^2_1=0.017; P_{FDR}=.99; PR=2.33), suggesting behavioral and psychiatric manifestations which are consistent with the FTD enrichment and earlier onset noted above. Notably, although these diagnoses did not reach significance in cluster 2, they were absent in the other clusters. Moreover, many diagnoses in cluster 2 were marked with a lack of PR, indicating a lower relevance of these diagnoses compared to clusters 1 or 3. Cluster 3 exhibited significant increases in diagnoses from a variety of categories, including respiratory issues (eg, cough: χ^2_1=339.47; P_{FDR}<.001; PR=3.77 and shortness of breath: χ^2_1=306.60; P_{FDR}<.001; PR=6.03), chronic pain (χ^2_1=460.55; P_{FDR}<.001; PR=4.80), musculoskeletal problems (eg, bilateral low back pain without sciatica: χ^2_1=456.91; P_{FDR}<.001; PR=4.80 and foot pain: χ^2_1=372.91; P_{FDR}<.001; PR=3.86), and complications of diabetes mellitus. Notably, diabetes mellitus (type 2 with autonomic neuropathy: χ^2_1=322.98; P_{FDR}<.001; PR=11.60 and insulin-dependent diabetes: χ^2_1=308.96; P_{FDR}<.001; PR=11.30) had exceptionally high PRs, suggesting a very strong association with these severe diabetes conditions in cluster 3.

Table 2. Summary statistics of International Classification of Diseases clusters.

Characteristics	Total (N=3454)	Cluster 1 (n=1501)	Cluster 2 (n=1597)	Cluster 3 (n=356)
Age of onset[a] (y), mean (SD)	72.1 (9.5)	73.6 (9.1)	69.8 (9.6)	75.5 (8.1)
Sex, n (%)				
Female	1678 (48.58)	709 (47.24)	769 (48.15)	200 (56.2)
Male	1776 (51.42)	792 (52.76)	828 (51.85)	156 (43.8)
Race, n (%)				
White	3059 (88.56)	1335 (88.94)	1421 (88.98)	303 (85.1)
Black or African American	77 (2.23)	37 (2.47)	25 (1.57)	15 (4.2)
Asian	90 (2.61)	40 (2.66)	40 (2.5)	10 (2.8)
American Indian or Alaska Native	4 (0.12)	0 (0)	2 (0.13)	2 (0.6)
Native Hawaiian or Other Pacific Islander	1 (0.03)	0 (0)	1 (0.06)	0 (0)
Other	103 (2.98)	51 (3.4)	34 (2.13)	18 (5.1)
Unavailable	120 (3.47)	38 (2.53)	74 (4.63)	8 (2.2)
Ethnicity, n (%)				
Not Hispanic or Latino	3020 (87.43)	1358 (90.47)	1326 (83.03)	336 (94.4)
Hispanic or Latino	106 (3.07)	45 (3)	42 (2.63)	19 (5.3)
Unavailable	328 (9.5)	98 (6.53)	229 (14.34)	1 (0.3)
ADRD Dx[b], n (%)				
AD[c]	1317 (38.13)	584 (38.91)	636 (39.82)	97 (27.2)
Dementia unspecified	1101 (31.88)	540 (35.98)	388 (24.3)	173 (48.6)
FTD[d]	190 (5.5)	64 (4.26)	121 (7.58)	5 (1.4)
LBD[e]	261 (7.56)	111 (7.4)	127 (7.95)	23 (6.5)
VCI[f]	96 (2.78)	44 (2.93)	33 (2.07)	19 (5.3)
Other[g]	489 (14.16)	158 (10.53)	292 (18.28)	39 (11)

[a]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

[b]Dx: diagnosis.

[c]AD: Alzheimer disease.

[d]FTD: frontotemporal dementia.

[e]LBD: Lewy body dementia.

[f]VCI: vascular cognitive impairment.

[g]Includes posterior cortical atrophy, progressive supranuclear palsy, corticobasal degeneration, and primary progressive aphasia.

Furthermore, statistical analyses revealed significant associations between ICD cluster membership, demographic variables, and diagnostic categories. Cluster 3 had an overrepresentation of female participants relative to clusters 1 and 2 (χ^2_1=8.8; P_{FDR}=.009; PR=1.178); however, clusters 1 and 2 showed no significant differences in sex distribution (cluster 1: χ^2_1=1.8; P_{FDR}=.26 and cluster 2: χ^2_1=0.2; P_{FDR}=.66; Figure 4B). In addition, the age of onset varied significantly among the ICD clusters (Kruskal-Wallis χ^2_2=182.6; P_{FDR}<.001, η^2=0.052). Cluster 2, with a mean age of onset of 69.8 (SD 9.6) years, had a significantly earlier age of onset compared with clusters 1 (Z=10.13; P_{FDR}<.001) and 3 (Z=−8.83; P_{FDR}<.001). Moreover, cluster 1 (mean 73.6, SD 9.1 years) had an earlier onset than cluster 3 (mean 75.5, SD 8.1 years; Z=−2.61; P_{FDR}=.009; Figure 4C). Cluster 1 was significantly enriched by dementia unspecified (χ^2_1=20.2; $_{FDR}$<.001; PR=1.252); cluster 2 was significantly enriched by FTD (χ^2_1=23.9; $_{FDR}$<.001; PR=2.039) and other rare ADRDs (χ^2_1=41; $_{FDR}$ <.001; PR=1.724); and cluster 3 was significantly enriched by VCI (χ^2_1=8.6; $_{FDR}$=.007; PR=2.147) and dementia unspecified (χ^2_1=50.2; $_{FDR}$<.001; PR=1.622). In contrast, no cluster was significantly enriched by AD, though AD was significantly underrepresented in cluster 3 (cluster 1: χ^2_1=0.6; $_{FDR}$=.53; cluster 2: χ^2_1=3.5; $_{FDR}$=.09; and cluster 3: χ^2_1=19.4; $_{FDR}$<.001; Figure 4D) or LBD (cluster 1: χ^2_1=0.1; $_{FDR}$=.80; cluster 2: χ^2_1=0.5; $_{FDR}$=.53; and cluster 3: χ^2_1=0.5; $_{FDR}$=.53; Figure 4D). Additional visualizations of

the UMAP projections colored by sex, age of onset, and ADRD diagnoses are available in Figures S3A, S3B, and S3C, respectively, in Multimedia Appendix 2.

Note Clustering

Initially, a provider effect was detected in the UMAP projection of note embeddings from the latest clinical notes of 3454 patients (Figure S4A in Multimedia Appendix 2). To address this, we used an optimal transport method, aligning the embeddings from all providers to the 2D embedding of a selected reference provider (Figure S4B in Multimedia Appendix 2). Following this alignment, hierarchical agglomerative clustering was applied to the adjusted note embeddings, revealing 3 distinct clusters, as determined by the silhouette score. The adjusted UMAP projection, color coded by cluster, is presented in Figure 5A. The patient distribution within these clusters was as follows: cluster 1 included 1280 (37.06%) patients, cluster 2 included 1161 (33.61%) patients, and cluster 3 included 1013 (29.33%) patients. Detailed summary statistics for each cluster are outlined in Table 3.

Figure 5. Clustering of note embeddings and their demographic and Alzheimer disease and related dementias (ADRD) diagnosis associations. (A) Uniform Manifold Approximation and Projection (UMAP) visualization of note embeddings characterized by 3 clusters: cluster 1 includes 1280 (37.1%) patients, cluster 2 includes 1161 (33.6%) patients, and cluster 3 includes 1013 (29.3%) patients. (B) Bar plot showing prevalence for sex by cluster, with significance based on chi-square tests. Notably, cluster 1 and cluster 3 were both enriched by female participants (P<.001) while cluster 2 was enriched by male participants (P<.001). (C) Violin plot illustrating the distribution of age of onset across clusters. Each violin plot shows the kernel density estimate of the data, with the center line representing the median age of onset. Box plot elements are overlaid, where the box limits indicate the upper and lower quartiles, and the whiskers extend to 1.5 times the IQR. Individual points are hidden for clarity. Significant differences are observed, with cluster 2 showing the latest average age of onset at 72.8 (SD 9.2) years, and cluster 1 the earliest at 71.5 (SD 9.3) years (P<.001). (D) Heatmap showing prevalence ratio for ADRD diagnoses across clusters, with significance derived from a chi-square test. Diagnoses include Alzheimer disease (AD); dementia unspecified; frontotemporal dementia (FTD); Lewy body dementia (LBD); vascular cognitive impairment (VCI); and others such as posterior cortical atrophy (PCA), progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), and primary progressive aphasia (PPA). No significant distribution variations are observed across clusters (P>.05). The circle size, color, and number indicate the magnitude of the prevalence ratio. Blue indicates underrepresentation while red indicates overrepresentation. *P* values were corrected with false discovery rate (FDR) adjustments.

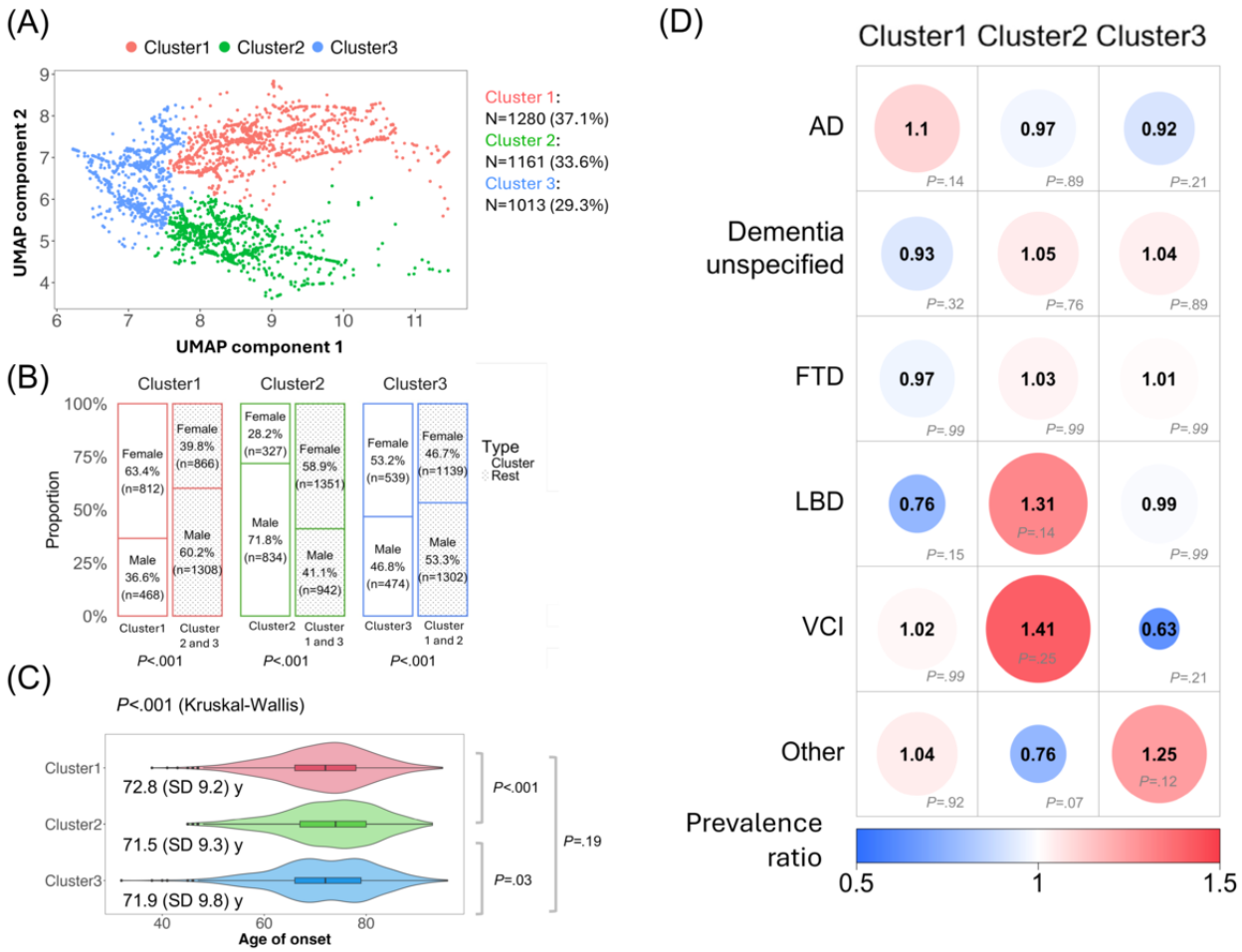


Table 3. Summary statistics of note clusters.

Characteristics	Total (N=3454)	Cluster 1 (n=1280)	Cluster 2 (n=1161)	Cluster 3 (n=1013)
Age of onset[a] (y), mean (SD)	72.1 (9.5)	71.5 (9.3)	72.8 (9.2)	71.9 (9.8)
Sex, n (%)				
Female	1678 (48.58)	812 (63.44)	327 (28.17)	539 (53.21)
Male	1776 (51.42)	468 (36.56)	834 (71.83)	474 (46.79)
Race, n (%)				
White	3059 (88.56)	1129 (88.20)	1032 (88.89)	898 (88.56)
Black or African American	77 (2.23)	30 (2.34)	26 (2.24)	21 (2.07)
Asian	90 (2.61)	36 (2.81)	25 (2.15)	29 (2.86)
American Indian or Alaska Native	4 (0.12)	1 (0.08)	1 (0.09)	2 (0.2)
Native Hawaiian or Other Pacific Islander	1 (0.03)	0 (0)	0 (0)	1 (0.1)
Other	103 (2.98)	34 (2.66)	44 (3.79)	25 (2.47)
Unavailable	120 (3.47)	50 (3.91)	33 (2.84)	37 (3.65)
Ethnicity, n (%)				
Not Hispanic or Latino	3020 (87.43)	1128 (88.13)	1004 (86.48)	888 (87.66)
Hispanic or Latino	106 (3.07)	36 (2.81)	39 (3.36)	31 (3.06)
Unavailable	328 (9.5)	116 (9.06)	118 (10.16)	94 (9.28)
ADRD Dx[b], n (%)				
AD[c]	1317 (38.13)	519 (40.55)	435 (37.47)	363 (35.83)
Dementia unspecified	1101 (31.88)	389 (30.39)	381 (32.82)	331 (32.68)
FTD[d]	190 (5.5)	69 (5.39)	65 (5.6)	56 (5.53)
LBD[e]	261 (7.56)	81 (6.33)	104 (8.96)	76 (7.5)
VCI[f]	96 (2.78)	36 (2.81)	40 (3.45)	20 (1.97)
Other[g]	489 (14.16)	186 (14.53)	136 (11.71)	167 (16.49)

[a]Age of onset was manually annotated by experts viewing clinical notes; for notes without age of onset, we approximated with the age of first encounter.

[b]Dx: diagnosis.

[c]AD: Alzheimer disease.

[d]FTD: frontotemporal dementia.

[e]LBD: Lewy body dementia.

[f]VCI: vascular cognitive impairment.

[g]Includes posterior.

We extracted common topics from each note cluster using topic modeling and examined the distribution of ADRD diagnoses across these clusters. In cluster 1, we found more terms related to psychiatric manifestations (eg, compulsive, indifference, and anxiety) and medications (eg, clozaril, trazodone, and sertraline), with a slight but nonsignificant enrichment in AD diagnosis (χ^2_1=4.9; P_{FDR}=.14; PR=1.10). Cluster 2 had more terms related to cardiovascular issues (eg, pacemaker, hypotension, and arrhythmia) and motor and sensory issues (eg, slurring, cochlear, and tremors), with a slight but nonsignificant enrichment in LBD (χ^2_1=4.6; P_{FDR}=.14; PR=1.31) and VCI (χ^2_1=2.5; P_{FDR}=.25; PR=1.41) diagnoses. Cluster 3 encompassed a wide variety of symptoms and conditions common in geriatric populations, including autoimmune (eg, scleroderma and vasculitis), behavioral changes and movement (eg, usual jocular behavior and imbalance, dysphagia), sleep (eg, insomnia), and sensory (eg, diplopia) problems, with a slight but not-significant enrichment in rare ADRD diagnoses (χ^2_1=3; P_{FDR}=.12; PR=1.25). Figure 3B depicts the list of representative words from each cluster and their PR, and Figure S5D in Multimedia Appendix 2 illustrates sentence examples from each cluster.

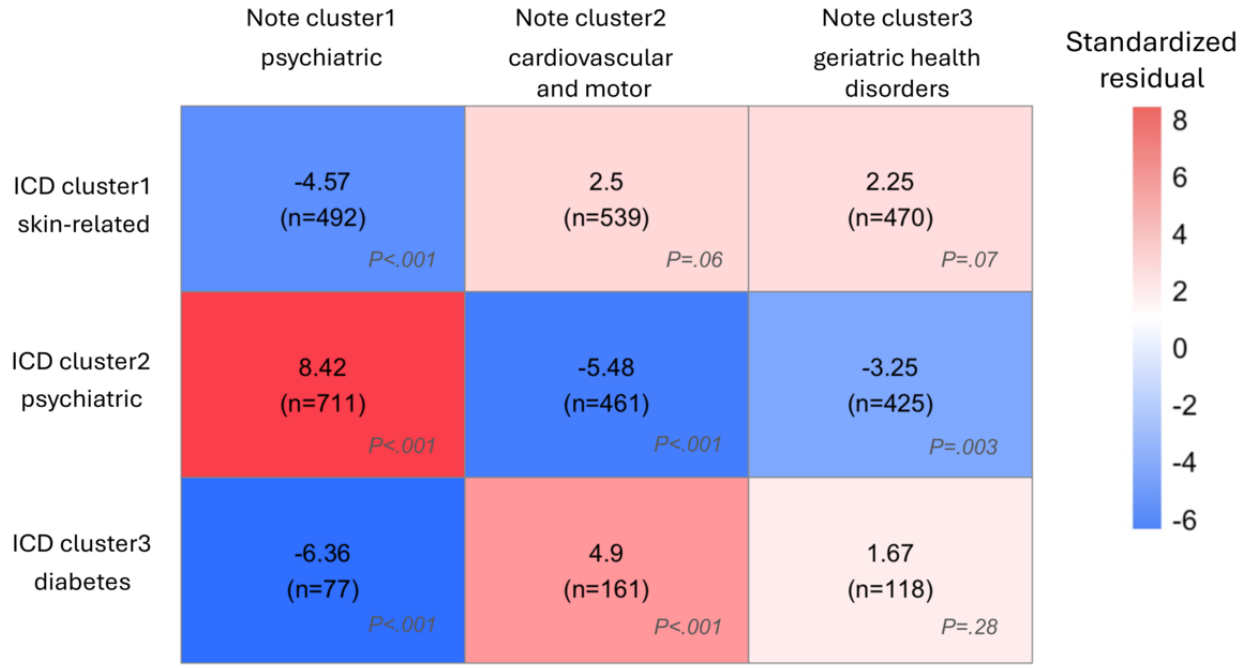
Furthermore, statistical analyses revealed significant associations of note cluster membership, with demographic variables, but not with ADRD diagnoses. First, both cluster 1 and 3 were significantly enriched by female participants (cluster 1: χ^2_1=178.7; P_{FDR}<.001; PR=1.593 and cluster 3: χ^2_1=12; P_{FDR}<.001; PR=1.14) while cluster 2 was enriched by male participants (χ^2_1=290.5; P_{FDR}<.001; PR=1.749; Figure 5B). In

addition, age of onset varied significantly among the note clusters (Kruskal-Wallis χ^2_2=14.9; P<.001; η^2=0.004), with cluster 2 (mean 72.8, SD 9.2 years) having a significantly later age of onset compared to cluster 1 (Z=−3.82; P_{FDR}<.001) and cluster 3 (Z=2.31; P_{FDR}=.03), while cluster 1 (mean 71.5, SD 9.3 years) and cluster 3 (mean 71.9, SD 9.8 years) did not differ (Z=−1.33; P_{FDR}=.19; Figure 5C). However, no association was observed between note cluster membership and ADRD diagnoses (Cluster 1-AD: χ^2_1=4.9; P_{FDR}=.14; Cluster 1-dementia unspecified: χ^2_1=1.9; P_{FDR}=.32; Cluster 1-FTD: χ^2_1=0.02; P_{FDR}=.99; Cluster 1-LBD: χ^2_1=4.1; P_{FDR}=.15; Cluster 1-other: χ^2_1=0.19; P_{FDR}=.92; Cluster 1-VCI: χ^2_1<0.001, P_{FDR}=.99; Cluster 2-AD: χ^2_1=0.3; P_{FDR}=.89; Cluster 2-dementia unspecified: χ^2_1=0.6; P_{FDR}=.76; Cluster 2-FTD: χ^2_1=0.01; P_{FDR}=.99; Cluster 2-LBD: χ^2_1=4.6; P_{FDR}=.14; Cluster 2-other: χ^2_1=8.3; P_{FDR}=.07; Cluster 2-VCI: χ^2_1=2.5; P_{FDR}=.25; Cluster 3-AD: χ^2_1=3.1; P_{FDR}=.21; Cluster 3-dementia unspecified: χ^2_1=0.4; P_{FDR}=.89; Cluster 3-FTD: χ^2_1<0.001; P_{FDR}=.99; Cluster 3-LBD: χ^2_1<0.001; P_{FDR}=.99; Cluster 3-other: χ^2_1=6.1; P_{FDR}=.12; Cluster 3-VCI: χ^2_1=3; P_{FDR}=.21; Figure 5D). Additional visualizations of the UMAP projections colored by sex, age of onset, and ADRD diagnoses are provided in Figures S5A, S5B, and S5C, respectively, in Multimedia Appendix 2.

Comparison Between ICD Clusters and Note Clusters

Statistical analysis demonstrated significant associations between *ICD* and note clusters (χ^2_4=89.43; P<.001; Cramér V=0.114). Specifically, note cluster 1, characterized by more female participants (PR=1.593) and terms related to psychiatric manifestations and medications, significantly overlapped (standardized residual=8.42; P_{FDR}<.001) with *ICD* cluster 2, which is noted for the earliest onset of disease (mean 69.8, SD 9.6 years) and a higher prevalence of psychiatric disorders and higher proportion of patients with FTD (PR=2.039). In addition, note cluster 2, which had higher proportion of male participants (PR=1.749) and terms related to cardiovascular and motor issues, overlapped significantly (standardized residual=4.90; P_{FDR}<.001) with *ICD* cluster 3, which is marked by the oldest onset of disease (mean 75.5, SD 8.1 years), a higher occurrence of VCI (PR=2.147), and dementia unspecified (PR=1.622) and had high prevalence of diabetes. These findings suggest a meaningful pattern of cluster correspondence across modalities (Figure 6).

Figure 6. Heatmap of the association of International Classification of Diseases (ICD) clusters with note clusters. Heatmap of the association of ICD clusters with note clusters. ICD clusters were significantly associated with note clusters (P<.001). Post hoc analyses revealed that note cluster 1 was positively associated with ICD cluster 2 (standardized residual=8.42, P<.001) and negatively associated with ICD clusters 1 and 3 (cluster 1: standardized residual=−4.57, P<.001; cluster 3: standardized residual=−6.36, P<.001). Furthermore, note cluster 2 was positively associated with ICD cluster 3 (standardized residual=4.9, P<.001) and negatively associated with ICD cluster 2 (standardized residual=−5.48, P<.001). Finally, note cluster 3 was negatively associated with ICD cluster 2 (standardized residual=−3.25, P=.003). Each cell displays standardized residuals (standardized differences between the observed count and the expected count) along with the count of overlapping patients (n). Color bar represents the value of the standardized residual. P values were corrected with false discovery rate adjustments.



Discussion

Principal Findings

This study aimed to characterize the clinical heterogeneity across ADRD by applying representation learning techniques on patient EHRs in a tertiary care clinic. We used pretrained *ICD* code embeddings, a transformer architecture for encoding clinical notes, and unsupervised learning to identify distinct ADRD subtypes. This work represents the first example of clustering patients with ADRD using embeddings derived directly from an LLM, without prior rule-based extraction of relevant medical concepts from the clinical note. The *ICD* codes allowed us to investigate subtypes of patients with similar *ICD* codes (non-ADRD diagnosis in charts), while the clinical notes allowed us to capture clinical history and presentation recorded by memory specialists. Our results demonstrate distinct patterns of disease manifestation with significant overlap between the *ICD* and note clusters. The overlap of clusters between the 2 approaches suggests that the subtypes may reflect common underlying clinical heterogeneity, as distinct subtypes can be identified through different data modalities.

Our choice of hierarchical agglomerative clustering was guided by the hierarchical nature of our data, empirical evidence from prior dementia subtyping studies, and theoretical limitations of alternative algorithms. Alternative methods, such as K-means, Gaussian mixture models [40], and density-based spatial clustering of applications with noise (DBSCAN) [41], face theoretical limitations: K-means assumes spherical clusters that is difficult to satisfy in high-dimensional embeddings; DBSCAN relies on density thresholds that break down in such spaces; and Gaussian mixture models can be unstable with overlapping subtypes. In contrast, hierarchical clustering preserves these nested relationships, resulting in more homogeneous [42] and reproducible [43] clusters.

In our study, we identified 3 clusters from *ICD* embeddings, each characterized by distinct health conditions. Cluster 1 predominantly featured issues related to skin health, with a commonality of dementia unspecified diagnoses and an average age of onset in the early 70s. The association between skin health and dementia in this cluster may be attributed to age, as both conditions become more prevalent with advancing age. This aligns with findings from previous studies, which reported an increase in the prevalence of actinic keratosis [44] and seborrheic keratosis [45], as well as AD [46], with age. Cluster 2 is marked by early-onset, psychiatric and behavioral manifestations; enrichment in FTD and other less common forms of ADRD; and the earliest age of onset among our clusters, typically in the late 60s. This cluster extends the characterization found in previous studies that described a behavioral symptom subtype in patients [20,21] by demonstrating similar characteristics in a broader population of patients with ADRD. Cluster 3 encompasses a broad array of conditions like respiratory issues and severe diabetes, affecting older patients, more female participants than male participants, with a higher incidence of VCI and dementia unspecified, and an age of onset in the mid-70s. Reflecting the AD subtype identified by previous researchers, this group was characterized as being overall older and having more comorbidities [20]. Landi et al [23] further differentiated patients with AD by onset timing, which we also observed but across a more diverse set of ADRD diagnoses. Notably, our *ICD*-based clustering did not reveal clearly separated clusters in the 2D UMAP projections. While UMAP aims to preserve both local and global relationships when reducing high-dimensional data to a lower-dimensional space, some distortions may inevitably occur during dimensionality reduction. Alternatively, the overlapping clusters could reflect the complexity of comorbid conditions in ADRD, which may not form clearly distinguishable subgroups.

In addition, our analysis of clinical notes revealed 3 distinct subtypes. Cluster 1 featured terms related to psychiatric manifestations and medications, aligning with findings from previous studies [20,21]. Cluster 2 included terms related to cardiovascular and various motor and sensory issues, supported by previous studies that identified subtypes of CVD [20,21] and aligning with the predominant diagnoses of VCI and LBD within this cluster. Cluster 3 covered a wide array of health conditions, consistent with the higher occurrence of rare ADRD diagnoses, which tend to involve more heterogeneous health conditions. Notably, we observed significant overlap between *ICD* and note clusters, identifying 2 ADRD subtypes of interest that were concordant across the 2 data modalities: the first subtype, “psychiatric manifestations,” and the second subtype, “diabetes with cardiovascular or motor issues.” Thus, our analysis delineated 2 distinct ADRD subtypes with specific diagnostic and symptomatic profiles.

Our study identified sex differences across all clinical note clusters with substantial effects observed in note clusters 1 and 2. For example, note cluster 1 was significantly overrepresented in female participants and had a higher prevalence of AD (PR=1.1). It also overlapped significantly (P<.001) with *ICD* cluster 2, which was enriched for psychiatric and behavioral symptoms (eg, apathy). This aligns with Tang et al [25], who reported stronger psychiatric associations in female patients with AD, including greater links to depression. This pattern may be partially attributed to women’s greater likelihood of seeking mental health care [47,48]. In contrast, note cluster 2 was overrepresented in male participants, with higher prevalence of VCI (PR=1.41) and LBD (PR=1.31), consistent with Tang et al [25], who found vascular dementia was more common in male participants. The higher prevalence of VCI in male participants may be related to a greater burden of hypertension, particularly in early life [49]. In addition, the increased representation of male participants with LBD may reflect potential underdiagnosis in female participants [50,51]. Given the clinical impact of sex disparities in ADRD—particularly in AD and LBD [52]—future studies integrating longitudinal data and clinicopathological evidence will be crucial to disentangling biological influences from health care–seeking behaviors.

Another interesting observation relates to variations in age of onset, a key indicator of disease severity, across the identified subtypes in both clinical notes and *ICD*-based clusters, with notable differences in the *ICD*-derived subtypes. For instance, the early-onset *ICD* cluster 2 was enriched with psychiatric disorders and included diseases known with early-onset,

including FTD [53] and other rare ADRD categories, such as PCA [54]. In contrast, the late-onset *ICD* cluster 3, exhibited a higher prevalence of diverse health conditions, including respiratory issues (eg, cough and shortness of breath), chronic pain, musculoskeletal conditions, such as bilateral low back pain and foot pain, as well as diabetes mellitus. While chronic pain is not typically associated with ADRD, pain could indicate the general aging process [55], and relate to osteoporosis and osteoarthritis, which likely contribute to chronic pain in older adults. Other symptoms, such as foot pain and shortness of breath, may reflect comorbidities of diabetes.

Limitations

Our study has a few limitations. First, there are the challenges associated with using real-world EHR data. Differences in how health care providers document information, stemming from variations in training, personal documentation habits, and clinical judgment, may have contributed to inconsistencies. Furthermore, health care use patterns, such as visit regularity, may influence our clustering results. For example, variations in visit frequency could lead to overrepresentation of certain symptom clusters or skewed associations. Future studies adjusting for health care use patterns may help address this limitation. Second, the inclusion of long-term patient histories in clinical notes—where recent notes may capture both current and past symptoms—could introduce extraneous information, making it difficult to isolate content relevant to the latest diagnosis. This mixture of historical and recent data may have diluted the association between documented ADRD diagnoses and their actual clinical significance, leading to observed trends rather than clear associations. Furthermore, the repetition of relevant language across multiple encounters may have influenced the clustering process, potentially reflecting the frequency of patient visits to the memory clinic, rather than clinical characteristics. Third, our study is constrained by the absence of an independent validation cohort to confirm the identified clusters. While the relative overlap in clusters identified through *ICD* codes and note contents, along with the alignment with findings from previous research, offers some validation, the results could be strengthened by applying the same encoding and clustering techniques to an external validation cohort. To ensure external validity, these results need to be validated at other health care institutions. Fourth, another limitation of this study is that our subtyping characterization only focused on ADRD diagnoses based on etiology, but did not address the stage of disease, which clearly affects the neuropsychological profile. This calls for a focus on the heterogeneity of disease stage in future research. Moreover, there may be sex differences in who receives health care at different stages and ages, adding another layer of complexity to our findings. Finally, our study is limited by the lack of racial diversity in the cohort, with 88.6% of participants being White. Given known racial differences in AD incidence, comorbidities, and health care access [56-59]—and particularly the heightened impact of hypertension on AD risk in some minoritized groups [59,60]—our findings may not fully capture the spectrum of ADRD subtypes in these populations. Future studies with broader representation are necessary to improve the generalizability of our subtyping approach.

Future Directions

In future work, the preprocessing of clinical notes could be enhanced by implementing multiple methods, such as medspaCy [61], with a focus on targeting sections most relevant to diagnoses, such as medical history. To further improve the extraction and analysis of pertinent data, the use of emerging LLMs, such as GPT [39], should be explored. In addition, validating an independent dataset and enriching the patient population would help increase the robustness and reliability of the identified ADRD subtypes. To advance this work, we will use a dual-modality approach that leverages both structured and unstructured data sources, such as medications and imaging. A deep autoencoder that uses multiple modalities simultaneously could offer methodological improvements over our current practice of conducting parallel clustering analyses and relying on heuristic averaging of embeddings. Furthermore, explicitly using the temporal or graph properties of EHRs could yield more informative representations, enhancing unsupervised clustering capabilities, as has been shown in prior approaches in a supervised learning setting [62,63]. Ultimately, our goal is to develop machine learning models capable of predicting these ADRD subtypes from real-word health care systems. Such models may aid in more precise diagnostics, prognostics, and the formulation of targeted treatment strategies.

Acknowledgments

This study was funded by National Institute of Aging P30AG062421 (SD and BTH). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. The authors acknowledge Yoonchul Shin and Qinglin Gou for annotating the age of onset from clinical notes.

Data Availability

The data used for this study are available from the Mass General Brigham Healthcare System, but restrictions apply to the availability of these data, which were used under permission for this study, and so are not publicly available. Code will be available on GitHub upon publication.

Authors' Contributions

MW was responsible for conceptualization, formal analysis, investigation, methodology, software, and writing original draft. YC was responsible for data curation, formal analysis, investigation, methodology, software, visualization, writing original draft,

review, and editing. YH was responsible for data curation and analysis, review, and editing. YL was responsible for formal analysis, methodology, and software. CM was responsible for conceptualization, review, and editing. BTH was responsible for review and editing. JRD was responsible for data curation, review, and editing. AS-P was responsible for review and editing. DB was responsible for review and editing. SD was responsible for conceptualization, funding acquisition, investigation, methodology, supervision, review, and editing.

Conflicts of Interest

JRD served on a scientific review board for I-Mab Biopharma. All other authors declare no conflicts of interest.

Multimedia Appendix 1
List of diagnosis names to Alzheimer disease and related dementias diagnosis categories. This document provides a list of diagnosis names from the electronic health records of patients at a memory clinic to various Alzheimer disease and related dementias diagnosis categories.
[PDF File (Adobe PDF File), 613 KB - aging_v8i1e65178_app1.pdf]

Multimedia Appendix 2
Visualizations of model attention and embedding representations. This appendix includes attention heatmaps from the final transformer layer across heads, as well as Uniform Manifold Approximation and Projection projections of International Classification of Diseases code, and clinical note embeddings, characterized by sex, age of onset, Alzheimer disease and related dementia diagnosis, and provider information.
[DOCX File , 2067 KB - aging_v8i1e65178_app2.docx]

References

1. Dementia fact sheet. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2024-04-29]
2. Alzheimer's disease fact sheet. National Institute on Aging. URL: https://www.nia.nih.gov/health/alzheimers-and-dementia/alzheimers-disease-fact-sheet [accessed 2024-04-29]
3. Bloom GS. Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol 2014 Apr 01;71(4):505-508. [doi: 10.1001/jamaneurol.2013.5847] [Medline: 24493463]
4. Stampfer MJ. Cardiovascular disease and Alzheimer's disease: common links. J Intern Med 2006 Sep 26;260(3):211-223 [FREE Full text] [doi: 10.1111/j.1365-2796.2006.01687.x] [Medline: 16918818]
5. Kotzbauer PT, Trojanowsk JQ, Lee VM. Lewy body pathology in Alzheimer's disease. J Mol Neurosci 2001 Oct;17(2):225-232. [doi: 10.1385/jmn:17:2:225] [Medline: 11816795]
6. Irwin DJ, Grossman M, Weintraub D, Hurtig HI, Duda JE, Xie SX, et al. Neuropathological and genetic correlates of survival and dementia onset in synucleinopathies: a retrospective analysis. Lancet Neurol 2017 Jan;16(1):55-65 [FREE Full text] [doi: 10.1016/S1474-4422(16)30291-5] [Medline: 27979356]
7. Ferreira D, Przybelski SA, Lesnick TG, Lemstra AW, Londos E, Blanc F, et al. β-Amyloid and tau biomarkers and clinical phenotype in dementia with Lewy bodies. Neurology 2020 Dec 15;95(24):e3257-e3268 [FREE Full text] [doi: 10.1212/WNL.0000000000010943] [Medline: 32989106]
8. Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 2002 Jul 19;297(5580):353-356. [doi: 10.1126/science.1072994] [Medline: 12130773]
9. Mukhopadhyay S, Banerjee D. A primer on the evolution of aducanumab: the first antibody approved for treatment of Alzheimer's disease. J Alzheimers Dis 2021;83(4):1537-1552. [doi: 10.3233/JAD-215065] [Medline: 34366359]
10. Neff RA, Wang M, Vatansever S, Guo L, Ming C, Wang Q, et al. Molecular subtyping of Alzheimer's disease using RNA sequencing data reveals novel mechanisms and targets. Sci Adv 2021 Jan 08;7(2):eabb5398 [FREE Full text] [doi: 10.1126/sciadv.abb5398] [Medline: 33523961]
11. Young AL, Marinescu RV, Oxtoby NP, Bocchetta M, Yong K, Firth NC, Genetic FTD Initiative (GENFI), Alzheimer's Disease Neuroimaging Initiative (ADNI). Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. Nat Commun 2018 Oct 15;9(1):4273 [FREE Full text] [doi: 10.1038/s41467-018-05892-0] [Medline: 30323170]
12. Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, Strandberg OT, Alzheimer's Disease Neuroimaging Initiative, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. Nat Med 2021 May;27(5):871-881 [FREE Full text] [doi: 10.1038/s41591-021-01309-6] [Medline: 33927414]
13. Scheltens NM, Tijms BM, Koene T, Barkhof F, Teunissen CE, Wolfsgruber S, et al. Cognitive subtypes of probable Alzheimer's disease robustly identified in four cohorts. Alzheimers Dement 2017 Nov 17;13(11):1226-1236 [FREE Full text] [doi: 10.1016/j.jalz.2017.03.002] [Medline: 28427934]

14. Libon DJ, Drabick DA, Giovannetti T, Price CC, Bondi MW, Eppig J, et al. Neuropsychological syndromes associated with Alzheimer's/vascular dementia: a latent class analysis. J Alzheimers Dis 2014 Sep 16;42(3):999-1014. [doi: 10.3233/JAD-132147] [Medline: 25024329]
15. Scheltens NM, Galindo-Garre F, Pijnenburg YA, van der Vlies AE, Smits LL, Koene T, et al. The identification of cognitive subtypes in Alzheimer's disease dementia using latent class analysis. J Neurol Neurosurg Psychiatry 2016 Mar 17;87(3):235-243. [doi: 10.1136/jnnp-2014-309582] [Medline: 25783437]
16. Alashwal H, El Halaby M, Crouse JJ, Abdalla A, Moustafa AA. The application of unsupervised clustering methods to Alzheimer's disease. Front Comput Neurosci 2019 May 24;13:31 [FREE Full text] [doi: 10.3389/fncom.2019.00031] [Medline: 31178711]
17. Doshi-Velez F, Ge Y, Kohane I. Comorbidity clusters in autism spectrum disorders: an electronic health record time-series analysis. Pediatrics 2014 Jan;133(1):e54-e63 [FREE Full text] [doi: 10.1542/peds.2013-0819] [Medline: 24323995]
18. Luo Y, Eran A, Palmer N, Avillach P, Levy-Moonshine A, Szolovits P, et al. A multidimensional precision medicine approach identifies an autism subtype characterized by dyslipidemia. Nat Med 2020 Sep 10;26(9):1375-1379. [doi: 10.1038/s41591-020-1007-0] [Medline: 32778826]
19. Zhang X, Chou J, Liang J, Xiao C, Zhao Y, Sarva H, et al. Data-driven subtyping of Parkinson's disease using longitudinal clinical records: a cohort study. Sci Rep 2019 Jan 28;9(1):797 [FREE Full text] [doi: 10.1038/s41598-018-37545-z] [Medline: 30692568]
20. Xu J, Wang F, Xu Z, Adekkanattu P, Brandt P, Jiang G, et al. Data-driven discovery of probable Alzheimer's disease and related dementia subphenotypes using electronic health records. Learn Health Syst 2020 Oct 10;4(4):e10246 [FREE Full text] [doi: 10.1002/lrh2.10246] [Medline: 33083543]
21. Alexander N, Alexander DC, Barkhof F, Denaxas S. Using unsupervised learning to identify clinical subtypes of Alzheimer's disease in electronic health records. Stud Health Technol Inform 2020 Jun 16;270:499-503. [doi: 10.3233/SHTI200210] [Medline: 32570434]
22. Alexander N, Alexander DC, Barkhof F, Denaxas S. Identifying and evaluating clinical subtypes of Alzheimer's disease in care electronic health records using unsupervised machine learning. BMC Med Inform Decis Mak 2021 Dec 08;21(1):343 [FREE Full text] [doi: 10.1186/s12911-021-01693-6] [Medline: 34879829]
23. Landi I, Glicksberg BS, Lee HC, Cherng S, Landi G, Danieletto M, et al. Deep representation learning of electronic health records to unlock patient stratification at scale. NPJ Digit Med 2020;3:96 [FREE Full text] [doi: 10.1038/s41746-020-0301-z] [Medline: 32699826]
24. He Z, Tian S, Erdengasileng A, Charness N, Bian J. Temporal subtyping of Alzheimer's disease using medical conditions preceding Alzheimer's disease onset in electronic health records. AMIA Jt Summits Transl Sci Proc 2022;2022:226-235 [FREE Full text] [Medline: 35854753]
25. Tang AS, Oskotsky T, Havaldar S, Mantyh WG, Bicak M, Solsberg CW, et al. Deep phenotyping of Alzheimer's disease leveraging electronic medical records identifies sex-specific clinical associations. Nat Commun 2022 Feb 03;13(1):675 [FREE Full text] [doi: 10.1038/s41467-022-28273-0] [Medline: 35115528]
26. Tang AS, Rankin KP, Cerono G, Miramontes S, Mills H, Roger J, et al. Leveraging electronic health records and knowledge networks for Alzheimer's disease prediction and sex-specific biological insights. Nat Aging 2024 Mar 21;4(3):379-395 [FREE Full text] [doi: 10.1038/s43587-024-00573-8] [Medline: 38383858]
27. Vaswani A, Brain G, Shazeer N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv Preprint posted online on June 12, 2017 [FREE Full text]
28. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019 Presented at: NAACL-HLT '19; June 2-7, 2019; Minneapolis, MN p. 4171-4186 URL: https://aclanthology.org/N19-1423.pdf
29. Alsentzer E, Murphy JR, Boag WH, Jindi D, Naumann T, McDermott M. Publicly available clinical BERT embeddings. In: Proceedings of the 2nd Clinical Natural Language Processing Workshop. 2019 Presented at: ClinicalNLP '19; June 7, 2019; Minneapolis, MN p. 72-78 URL: https://aclanthology.org/W19-1909.pdf [doi: 10.18653/v1/w19-1909]
30. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035 [FREE Full text] [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
31. Choi Y, Chiu CY, Sontag D. Learning low-dimensional representations of medical concepts. AMIA Jt Summits Transl Sci Proc 2016;2016:41-50 [FREE Full text] [Medline: 27570647]
32. Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27:63. [doi: 10.1002/j.1538-7305.1948.tb00917.x]
33. Ebrahimi N, Pflughoeft K, Soofi ES. Two measures of sample entropy. Stat Probabil Lett 1994 Jun;20(3):225-234. [doi: 10.1016/0167-7152(94)90046-9]
34. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar 3;17(3):261-272 [FREE Full text] [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]

35. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14 [FREE Full text] [doi: 10.3389/fninf.2014.00014] [Medline: 24600388]
36. Flamary R, Courty N. POT python optimal transport library. GitHub. URL: https://pythonot.github.io/ [accessed 2024-04-29]
37. Courty N, Flamary R, Tuia D, Rakotomamonjy A. Optimal Transport for Domain Adaptation. IEEE Trans Pattern Anal Mach Intell 2017 Sep 1;39(9):1853-1865. [doi: 10.1109/tpami.2016.2615921]
38. Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022 [FREE Full text]
39. Achiam OJ, Adler S, Agarwal S, Ahmad L, Akkaya I, Aleman FL, et al. GPT-4 technical report. arXiv Preprint posted online on March 15, 2023 [FREE Full text]
40. Reynolds D. Gaussian mixture models. In: Li SZ, Jain A, editors. Encyclopedia of Biometrics. Cham, Switzerland: Springer; 2009:659-663.
41. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. 1996 Presented at: KDD '96; August 2-4, 1996; Portland, OR p. 226-231 URL: https://dl.acm.org/doi/10.5555/3001460.3001507
42. Ribino P, Napoli CD, Paragliola G, Serino L, Gasparini F, Chicco D. Exploratory analysis of longitudinal data of patients with dementia through unsupervised techniques. In: Proceedings of the 4th Italian Workshop on Artificial Intelligence for an Ageing Society. 2023 Presented at: AIxAS '23; November 6-9, 2023; Rome, Italy p. 1-21 URL: https://ceur-ws.org/Vol-3623/AIxAS_2023_paper_8.pdf
43. Meng W, Xu J, Huang Y, Wang C, Song Q, Ma A, et al. Autoencoder to identify sex-specific sub-phenotypes in Alzheimer's disease progression using longitudinal electronic health records. medRxiv Preprint posted online on July 11, 2024 [FREE Full text] [doi: 10.1101/2024.07.07.24310055] [Medline: 39040206]
44. Yaldiz M. Prevalence of actinic keratosis in patients attending the dermatology outpatient clinic. Medicine (Baltimore) 2019 Jul;98(28):e16465 [FREE Full text] [doi: 10.1097/MD.0000000000016465] [Medline: 31305480]
45. Sun MD, Halpern AC. advances in the etiology, detection, and clinical management of seborrheic keratoses. Dermatology 2022 Jul 26;238(2):205-217. [doi: 10.1159/000517070] [Medline: 34311463]
46. Alzheimer's disease facts and figures. Alzheimer's Association. URL: https://www.alz.org/alzheimers-dementia/facts-figures?gad_source=1 [accessed 2024-04-29]
47. Lim MT, Fong Lim YM, Tong SF, Sivasampu S. Age, sex and primary care setting differences in patients' perception of community healthcare seeking behaviour towards health services. PLoS One 2019 Oct 21;14(10):e0224260 [FREE Full text] [doi: 10.1371/journal.pone.0224260] [Medline: 31634373]
48. Susukida R, Mojtabai R, Mendelson T. Sex differences in help seeking for mood and anxiety disorders in the national comorbidity survey-replication. Depress Anxiety 2015 Nov 22;32(11):853-860. [doi: 10.1002/da.22366] [Medline: 25903117]
49. Connelly PJ, Currie G, Delles C. Sex Differences in the prevalence, outcomes and management of hypertension. Curr Hypertens Rep 2022 Jun 07;24(6):185-192 [FREE Full text] [doi: 10.1007/s11906-022-01183-8] [Medline: 35254589]
50. Bayram E, Coughlin DG, Rajmohan R, Litvan I. Sex differences for clinical correlates of substantia nigra neuron loss in people with Lewy body pathology. Biol Sex Differ 2024 Jan 19;15(1):8 [FREE Full text] [doi: 10.1186/s13293-024-00583-6] [Medline: 38243325]
51. Bayram E, Coughlin DG, Koga S, Ross OA, Litvan I, Dickson DW. Sex differences for regional pathology in people with a high likelihood of Lewy body dementia phenotype based on underlying pathology. Alzheimers Dement (Amst) 2025;17(1):e70083. [doi: 10.1002/dad2.70083] [Medline: 39886324]
52. Wei H, Masurkar AV, Razavian N. On gaps of clinical diagnosis of dementia subtypes: a study of Alzheimer's disease and Lewy body disease. Front Aging Neurosci 2023 Mar 21;15:1149036 [FREE Full text] [doi: 10.3389/fnagi.2023.1149036] [Medline: 37025965]
53. Lee SE. Frontotemporal dementia: clinical features and diagnosis. UpToDate. URL: https://www.uptodate.com/contents/frontotemporal-dementia-clinical-features-and-diagnosis [accessed 2024-04-29]
54. Chapleau M, La Joie R, Yong K, Agosta F, Allen IE, Apostolova L, et al. Demographic, clinical, biomarker, and neuropathological correlates of posterior cortical atrophy: an international cohort study and individual participant data meta-analysis. Lancet Neurol 2024 Feb;23(2):168-177 [FREE Full text] [doi: 10.1016/S1474-4422(23)00414-3] [Medline: 38267189]
55. Cheng Y, Ho E, Weintraub S, Rentz D, Gershon R, Das S, et al. Predicting brain amyloid status using the National Institute of Health Toolbox (NIHTB) for assessment of neurological and behavioral function. J Prev Alzheimers Dis 2024;11(4):943-957. [doi: 10.14283/jpad.2024.77] [Medline: 39044505]
56. Adkins-Jackson PB, Kraal AZ, Hill-Jarrett TG, George KM, Deters KD, Besser LM, et al. Riding the merry-go-round of racial disparities in ADRD research. Alzheimers Dement 2023 Oct;19(10):4735-4742. [doi: 10.1002/alz.13359] [Medline: 37394968]
57. Mielke MM, Aggarwal NT, Vila-Castelar C, Agarwal P, Arenaza-Urquijo EM, Brett B, et al. Consideration of sex and gender in Alzheimer's disease and related disorders from a global perspective. Alzheimer's & Dementia 2022 Dec 08;18(12):2707-2724 [FREE Full text] [doi: 10.1002/alz.12662] [Medline: 35394117]

58. Hernandez S, McClendon MJ, Zhou XH, Sachs M, Lerner AJ. Pharmacological treatment of Alzheimer's disease: effect of race and demographic variables. J Alzheimers Dis 2010 Jan 07;19(2):665-672 [FREE Full text] [doi: 10.3233/JAD-2010-1269] [Medline: 20110610]
59. Steenland K, Tan Y, Wingo T, Shi L, Xiao S, Wharton W. The effect of race and co-morbidities on Alzheimer's disease based on Medicare data. Alzheimers Dement 2023 May;19(5):1858-1864 [FREE Full text] [doi: 10.1002/alz.12838] [Medline: 36327171]
60. Akushevich I, Kolpakov S, Yashkin AP, Kravchenko J. Vulnerability to hypertension is a major determinant of racial disparities in Alzheimer's disease risk. Am J Hypertens 2022 Aug 01;35(8):745-751 [FREE Full text] [doi: 10.1093/ajh/hpac063] [Medline: 35581146]
61. Eyre H, Chapman AB, Peterson KS, Shi J, Alba PR, Jones MM, et al. Launching into clinical space with medspaCy: a new clinical text processing toolkit in Python. AMIA Annu Symp Proc 2021;2021:438-447 [FREE Full text] [Medline: 35308962]
62. Choi E, Xu Z, Li Y, Dusenberry M, Flores G, Xue E, et al. Learning the graphical structure of electronic health records with graph convolutional transformer. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence. 2020 Apr 03 Presented at: AAAI '20; February 7-12, 2020; New York, NY p. 606-613. [doi: 10.1609/aaai.v34i01.5400]
63. Zhu W, Razavian N. Variationally regularized graph-based representation learning for electronic health records. In: Proceedings of the 2021 Conference on Health, Inference, and Learning. 2021 Presented at: CHIL '21; April 8-10, 2021; Virtual Event p. 1-13 URL: https://dl.acm.org/doi/10.1145/3450439.3451855 [doi: 10.1145/3450439.3451855]

Abbreviations

AD: Alzheimer disease
ADRD: Alzheimer disease and related dementia
Aβ: amyloid beta
BERT: Bidirectional Encoder Representations from Transformers
CVD: cerebrovascular disease
DBSCAN: density-based spatial clustering of applications with noise
EHR: electronic health record
FDR: false discovery rate
FTD: frontotemporal dementia
ICD: International Classification of Diseases
LBD: Lewy body dementia
LLM: large language model
MGH: Massachusetts General Hospital
MIMIC: Medical Information Mart for Intensive Care
PCA: posterior cortical atrophy
PR: prevalence ratio
TF-IDF: term frequency–inverse document frequency
UMAP: Uniform Manifold Approximation and Projection
VCI: vascular cognitive impairment

Edited by D Liu; submitted 07.08.24; peer-reviewed by R Zapata, D Chrimes, A Chaturvedi; comments to author 29.01.25; revised version received 19.02.25; accepted 10.03.25; published 31.03.25.

Please cite as:
West M, Cheng Y, He Y, Leng Y, Magdamo C, Hyman BT, Dickson JR, Serrano-Pozo A, Blacker D, Das S
Unsupervised Deep Learning of Electronic Health Records to Characterize Heterogeneity Across Alzheimer Disease and Related Dementias: Cross-Sectional Study
JMIR Aging 2025;8:e65178
URL: https://aging.jmir.org/2025/1/e65178
doi:10.2196/65178
PMID:

Correction: Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

Natthanaphop Isaradech[1,2*], MD; Wachiranun Sirikul[1,3,4*], MD, MSc; Nida Buawangpong[5], MD; Penprapa Siviroj[1], DrPH; Amornphat Kitro[1,2], MD, MCTM

[1]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

[2]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Center of Data Analytics and Knowledge Synthesis for Health Care, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[*]these authors contributed equally

Corresponding Author:
Wachiranun Sirikul, MD, MSc
Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

Related Article:

Correction of: https://aging.jmir.org/2025/1/e62942

(JMIR Aging 2025;8:e75690) doi:10.2196/75690

KEYWORDS

aged care; gerontology; geriatric; old; aging; clinical decision support; delivering health information and knowledge to the public; diagnostic systems; digital health; epidemiology; surveillance; diagnosis; frailty; machine learning; prediction; predictive; AI; artificial intelligence; Thailand; community dwelling; health care intervention; patient care

In “Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study” (JMIR Aging 2025;8:e62942) one error was noted.

Reference 44 was a duplicate of reference 36, which reads as follows:

Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in northern Thailand: a cross-sectional study. Int J Environ Res Public Health. Jun 8, 2020;17(11):4077.

Reference 44 has therefore been removed, and all subsequent references have been reordered accordingly.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.04.25; this is a non–peer-reviewed article;accepted 10.04.25; published 22.04.25.

Please cite as:
Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A
Correction: Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study
JMIR Aging 2025;8:e75690
URL: https://aging.jmir.org/2025/1/e75690
doi:10.2196/75690

Correction: Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial

Yanhong Xie[1*], MSc; Shanshan Shen[1*], MSc; Caixia Liu[2], BSc; Hong Hong[3], BSc; Huilan Guan[1], BSc; Jingmei Zhang[1], BSc; Wanqi Yu[4], BSc

[1]Geriatrics Department, Zhejiang Hospital, Hangzhou, China

[2]Nursing Department, Zhejiang Hospital, Hangzhou, China

[3]General Surgery, Zhejiang Hospital, Hangzhou, Zhejiang Province, China

[4]The Medical Record Statistics Department, Zhejiang Hospital, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Hong Hong, BSc
General Surgery, Zhejiang Hospital, Hangzhou, Zhejiang Province, China

Related Article:

Correction of: https://aging.jmir.org/2024/1/e50847

(JMIR Aging 2025;8:e69493) doi:10.2196/69493

In "Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial" (JMIR Aging. 2024 Oct 4:7:e50847. doi: 10.2196/50847) the authors made three corrections.

The following funding information has been added to the end of the "Acknowledgements" section:

> *This work was supported by grants from the National Key Research and Development Program of China (program number: 2019YFE0113100), the Zhejiang Provincial Medicine and Health Technology Project (grant number: 2019KY003), and the Zhejiang Hospital 3060 Excellent Young Talents Training Project.*

Additionally, the authors would like to specify equal contribution between Yanhong Xie and Shanshan Shen as co-first authors, and Hong Hong and Caixia Liu as co-corresponding authors. The latter will be specified at the beginning of the "Acknowledgements" section as follows:

> *CL is the co-corresponding author on this work, and can be reached at the following email address: zjyyhlb2007@126.com.*

Finally, the following ORCID has been added to author Hong Hong:

> *0009-0001-1372-1670*

The corrections will appear in the online version of the paper on the JMIR Publications website on January 29, 2025, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 01.12.24; this is a non–peer-reviewed article;accepted 04.12.24; published 29.01.25.

Please cite as:
Xie Y, Shen S, Liu C, Hong H, Guan H, Zhang J, Yu W
Correction: Internet-Based Supportive Interventions for Family Caregivers of People With Dementia: Randomized Controlled Trial
JMIR Aging 2025;8:e69493
URL: https://aging.jmir.org/2025/1/e69493
doi:10.2196/69493

Association Between Sleep Duration and Cognitive Frailty in Older Chinese Adults: Prospective Cohort Study

Ruixue Cai[1,2], PhD; Jianqian Chao[1], PhD; Chenlu Gao[2,3], PhD; Lei Gao[2,3,4,5], MBBS; Kun Hu[2,3,4,5], PhD; Peng Li[2,3,4,5], PhD

[1]Department of Epidemiology and Health Statistics, School of Public Health, Southeast University, Nanjing, China

[2]Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, 149 13th Street, Boston, MA, United States

[3]Division of Sleep Medicine, Harvard Medical School, Boston, MA, United States

[4]Division of Sleep and Circadian Disorders, Departments of Neurology and Medicine, Brigham and Women's Hospital, Boston, MA, United States

[5]Broad Institute of MIT and Harvard, Cambridge, MA, United States

Corresponding Author:
Peng Li, PhD
Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, 149 13th Street, Boston, MA, United States

Abstract

Background: Disturbed sleep patterns are common among older adults and may contribute to cognitive and physical declines. However, evidence for the relationship between sleep duration and cognitive frailty, a concept combining physical frailty and cognitive impairment in older adults, is lacking.

Objective: This study aimed to examine the associations of sleep duration and its changes with cognitive frailty.

Methods: We analyzed data from the 2008 - 2018 waves of the Chinese Longitudinal Healthy Longevity Survey. Cognitive frailty was rendered based on the modified Fried frailty phenotype and Mini-Mental State Examination. Sleep duration was categorized as short (<6 h), moderate (6 - 9 h), and long (>9 h). We examined the association of sleep duration with cognitive frailty status at baseline using logistic regressions and with the future incidence of cognitive frailty using Cox proportional hazards models. Restricted cubic splines were used to explore potential nonlinear associations.

Results: Among 11,303 participants, 1298 (11.5%) had cognitive frailty at baseline. Compared to participants who had moderate sleep duration, the odds of having cognitive frailty were higher in those with long sleep duration (odds ratio 1.71, 95% CI 1.48 - 1.97; *P*<.001). A J-shaped association between sleep duration and cognitive frailty was also observed (*P*<.001). Additionally, during a mean follow-up of 6.7 (SD 2.6) years among 5201 participants who were not cognitively frail at baseline, 521 (10%) participants developed cognitive frailty. A higher risk of cognitive frailty was observed in participants with long sleep duration (hazard ratio 1.32, 95% CI 1.07 - 1.62; *P*=.008).

Conclusions: Long sleep duration was associated with cognitive frailly in older Chinese adults. These findings provide insights into the relationship between sleep duration and cognitive frailty, with potential implications for public health policies and clinical practice.

(JMIR Aging 2025;8:e65183) doi:10.2196/65183

KEYWORDS

aging; frailty; cognition; cohort study; sleep duration; sleep quality; longitudinal study

Introduction

Physical frailty and cognitive impairment are prevalent among older adults and have individually been associated with adverse health outcomes [1,2]. They often coincide with aging and can be bidirectionally linked to each other [3,4], prompting the introduction of the concept of cognitive frailty—the coexistence of both physical frailty and cognitive impairment [5]. The necessity is further justified by the findings that cognitive frailty poses an even greater risk of adverse outcomes compared to the isolated effects of the 2 conditions [6,7]. Identifying factors associated with or contributing to cognitive frailty is important and will provide a better understanding of the underlying mechanisms.

Disturbed sleep patterns, such as an increased number of awakenings, abnormal nighttime sleep duration, and poorer sleep quality, are common among older adults [8]. Numerous epidemiological studies identified sleep disturbances as risk factors for physical frailty and cognitive impairment, such as shorter or longer sleep duration, insomnia, and sleep apnea [9,10]. In addition, recent studies have linked long sleep duration to cognitive frailty, but they are from cross-sectional analyses

[11,12]. A longitudinal study in Mexican older adults reported associations between increasing sleep duration trajectory with greater odds for mild cognitive impairment and frailty separately [13]. Whether sleep patterns are associated with the development of cognitive frailty as a unified concept remains to be elucidated.

To address this knowledge gap, this study aimed to assess the associations of sleep duration with cognitive frailty in a nationally representative, large cohort of older Chinese adults. By focusing on an older Chinese adult population, our findings may offer insights into the relationship between sleep duration and cognitive frailty within this demographic, with potential implications for public health policies and clinical practice.

Methods

Study Design and Participants

We used data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). As detailed previously, CLHLS is an ongoing, nationwide representative longitudinal cohort study of Chinese adults aged 65 years and older from 23 out of 31 provinces or municipalities or autonomous regions in mainland China [14,15]. The CLHLS started in 1998 with surveys on sociodemographic characteristics, lifestyle, cognitive function, psychological status, and physical function conducted every 2 - 3 years.

In this study, we used data from 4 waves (2008, 2011, 2014, and 2018) of the CLHLS. To maximize statistical power and leverage the large sample size, we first examined cross-sectional associations between sleep duration and sleep quality, with cognitive frailty at baseline (ie, the 2008 wave). For this set of analyses, we included 11,303 participants who had cognition and physical frailty assessments at baseline. We then examined whether baseline sleep duration and sleep quality, as well as change in sleep duration, were associated with the future development of cognitive frailty. For these longitudinal analyses, we excluded 1298 participants who had cognitive frailty at baseline, 4780 participants who had no follow-up cognitive frailty assessment, and 24 participants who had missing value in sleep duration in the 2011 wave. As a result, 5201 participants entered the subsequent longitudinal analysis.

Assessment of Cognitive Frailty

Cognitive frailty was identified as the simultaneous presence of both physical frailty and cognitive impairment [5].

Physical frailty was assessed by a modified Fried frailty phenotype, based on 5 components: shrinking, weakness, low mobility, exhaustion, and inactivity [16,17]. Shrinking was defined as having a BMI of 18.5 kg/m^2 or less. Weakness was determined if participants self-reported that they were unable to lift a bag weighing 5 kg or above. Low mobility was identified if participants self-reported that they were unable to walk 1 km or longer in a row. Exhaustion was identified if participants answered “always,” “often,” or “sometimes” to the question “Do you feel the older you get, the more useless you are?” Inactivity was defined as participants doing the following activities 1 time per week or less: housework, outside activity, garden work, raising domestic animals or pets, playing cards or mah-jongg, and social activity (organized). Participants were considered physically frail when meeting ≥3 of the 5 criteria.

Cognitive impairment was assessed by a validated Chinese version of the Mini-Mental State Examination (C-MMSE) [15,18]. The C-MMSE includes 24 items regarding orientation, memory, attention, calculation, and language, ranging from 0 to 30. Based on a prior study, participants were considered cognitive impairment if the C-MMSE score was ≤22 [15].

Assessment of Sleep Variables

Sleep duration was estimated based on the answers to the question “How many hours do you sleep normally?” Based on their answers, participants were further categorized into short (<6 h), moderate (6 - 9 h), and long (>9 h) sleep durations. This categorization aligned with previous research indicating that both short and long sleep durations are associated with poorer health outcomes [19]. The rate of change in sleep duration was calculated by taking the difference between sleep durations reported in the 2008 and 2011 waves (ie, sleep duration in the 2011 wave minus sleep duration in the 2008 wave) and dividing it by the individual follow-up interval in years. Sleep quality was assessed based on answers to the question “How about the quality of your sleep?” A score of “poor” was assigned if the answer was “so-so,” “bad,” or “very bad,” and a score of “good” was assigned otherwise.

Covariates

We considered the following covariates that have been identified to be associated with physical or cognitive outcomes in previous studies [20,21]: age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity. Education was categorized as “Not Educated” if the participant had not attended any schools and “With Formal Education” if the participant had attended schools more than 1 year. Marital status was defined as currently married and living with a spouse, and others (eg, separated, divorced, widowed, and never married). Residence was categorized into living in urban (eg, city or town) or rural areas. Economic status was categorized into economic dependence or independence based on their income source. Loneliness was identified if participants answered “always,” “often,” or “sometimes” to the question “Do you feel lonely and isolated?” Smoking status was categorized as “Never smoked,” and “Former or current smoker.” Drinking status was categorized into “Never drank,” and “Former or current drinker.” Multimorbidity was defined from self-reports and categorized as “Yes” if having 2 or more of the following chronic diseases: hypertension, diabetes, heart disease, stroke or cerebrovascular disease, respiratory disease, cancer, and Parkinson disease, or “No” if otherwise.

Statistical Analysis

Listwise deletion was applied to both cross-sectional and longitudinal analyses to handle missing data on covariates. The *t* tests for continuous variables and chi-square tests for categorical variables were used to compare the differences in the baseline characteristics.

Multivariate logistic regression models were performed to assess the associations of sleep duration and sleep quality with

cognitive frailty at baseline (ie, the 2008 wave). We performed 3 models: model A was unadjusted; model B was adjusted for age, sex, and education; model C was further adjusted for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity.

Cox proportional hazard (PH) models were used to test the associations between sleep duration and sleep quality with incident cognitive frailty during the follow up. Death was treated as a right-censored event at the time of occurrence. The follow-up duration was defined as the time interval in years from baseline to the date of the first occurrence of cognitive frailty, death, or the latest available data, whichever occurred first. We rounded the time to the nearest integer year to account for group-tied events. Similarly, we performed 3 models: model A was unadjusted; model B was adjusted for age, sex, and education; model C was further adjusted for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity. The PH assumption of the Cox models was assessed using Schoenfeld residuals. No major violations of the PH assumption were detected for exposure variables.

Restricted cubic spline (RCS) curves were plotted to explore the nonlinear cross-sectional and longitudinal associations between sleep duration and cognitive frailty. Additionally, several sensitivity analyses were conducted to verify the robustness of our findings. First, we adjusted for each chronic disease individually and for all chronic diseases simultaneously. Second, we performed stratified analyses by sex and age. Furthermore, we performed similar sets of Cox PH models to investigate the associations of changes in sleep duration and incident cognitive frailty.

All statistical analyses were performed in R (version 4.1.2; R Core Team). Statistical significance was considered at a 2-tailed α level of .05.

Ethical Considerations

The CLHLS was approved by the Duke University Institutional Review Board (Pro00062871) and the Peking University Biomedical Ethics Committee (IRB00001052–13074). All participants provided written informed consent. The data used in this study were deidentified to protect participant privacy and confidentiality. No compensation was provided to participants.

Results

Participant Characteristics

The participant flowchart is presented in Figure 1. Of the 11,303 participants included in the cross-sectional analysis, 6037 (53.4%) participants were female. The mean age at baseline was 84.7 (SD 11.0) years old. Approximately 11.5% (n=1298) of all participants had cognitive frailty at baseline. The numbers of participants who reported short, moderate, or long sleep duration were 1327 (11.7%), 7207 (63.8%), and 2769 (24.5%), respectively. Poor sleep quality was found in 3796 (33.6%) participants. Baseline characteristics stratified by cognitive frailty status at baseline are summarized in Table 1. Additionally, baseline characteristics categorized by sleep duration are presented in Multimedia Appendix 1. Baseline characteristics of 5201 participants included in the longitudinal analysis are summarized in Multimedia Appendix 2.

Figure 1. Flowchart of participants through the study. CLHLS: Chinese Longitudinal Healthy Longevity Survey.

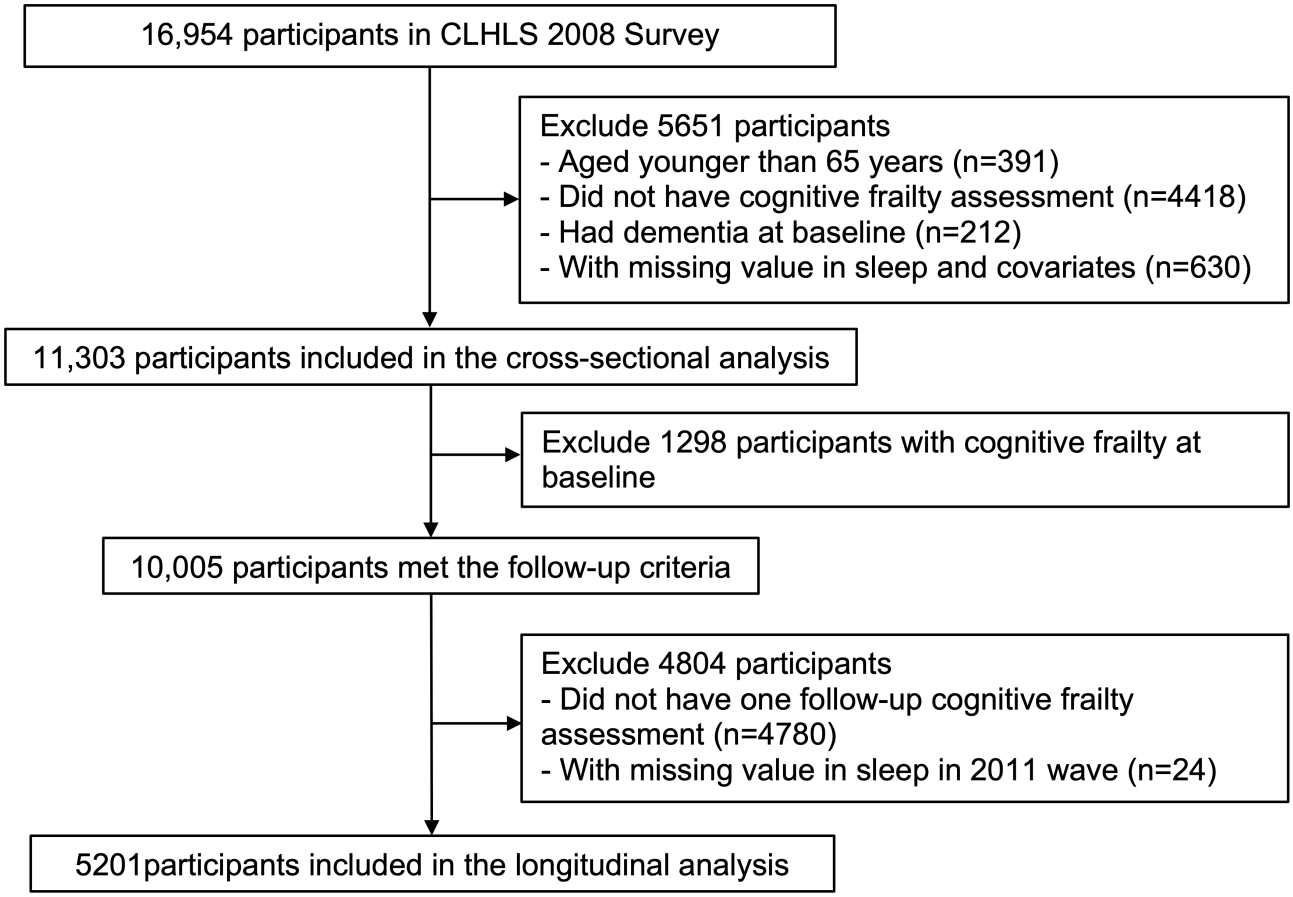


Table . Baseline characteristics of participants by cognitive frailty status at baseline.

	Overall (n=11,303)	Non-CF[a] (n=10,005)	CF (n=1298)	*P* value
Age (years), mean (SD)	84.7 (11.0)	83.3 (10.6)	95.3 (7.2)	<.001
Sex, n (%)				<.001
Male	5266 (46.6)	4966 (49.6)	300 (23.1)	
Female	6037 (53.4)	5039 (50.4)	998 (76.9)	
Education, n (%)				<.001
Not educated	6548 (57.9)	5457 (54.5)	1091 (84.1)	
With formal education	4755 (42.1)	4548 (45.5)	207 (15.9)	
Marital status, n (%)				<.001
Married and living with spouse	4098 (36.3)	3979 (39.8)	119 (9.2)	
Others	7205 (63.7)	6026 (60.2)	1179 (90.8)	
Current residence, n (%)				.12
Urban	4647 (41.1)	4140 (41.4)	507 (39.1)	
Rural	6656 (58.9)	5865 (58.6)	791 (60.9)	
Economic status, n (%)				<.001
Dependence	7929 (70.2)	6716 (67.1)	1213 (93.5)	
Independence	3374 (29.9)	3289 (32.9)	85 (6.5)	
Loneliness, n (%)				<.001
Yes	3581 (31.7)	2943 (29.4)	638 (49.2)	
No	7722 (68.3)	7062 (70.6)	660 (50.8)	
Smoking status, n (%)				<.001
Never smoked	7271 (64.3)	6230 (62.3)	1041 (80.2)	
Former or current smoker	4032 (35.7)	3775 (37.7)	257 (19.8)	
Drinking status, n (%)				<.001
Never drank	7670 (67.9)	6665 (66.6)	1005 (77.4)	
Former or current drinker	3633 (32.1)	3340 (33.4)	293 (22.6)	
Multimorbidity, n (%)				.02
Yes	1068 (9.4)	969 (9.7)	99 (7.6)	
No	10,235 (90.6)	9036 (90.3)	1199 (92.4)	
Sleep quality, n (%)				<.001
Good	7507 (66.4)	6706 (67.0)	801 (61.7)	
Poor	3796 (33.6)	3299 (33.0)	497 (38.3)	
Sleep duration, n (%)				<.001
Short (<6 h)	1327 (11.7)	1179 (11.8)	148 (11.4)	
Moderate (6 - 9 h)	7207 (63.8)	6557 (65.5)	650 (50.1)	
Long (>9 h)	2769 (24.5)	2269 (22.7)	500 (38.5)	

[a]CF: cognitive frailty.

Association Between Sleep Duration and Cognitive Frailty at Baseline

Compared to participants who had moderate sleep duration, the odds of having cognitive frailty were higher in those who had long sleep duration (odds ratio [OR] 1.71, 95% CI 1.48 - 1.97; *P*<.001) in the initial model adjusted for age, sex, and education Table 2. The association remained statistically significant with a similar OR in the augmented model after further adjustments of marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (OR 1.69, 95% CI 1.46 - 1.95; *P*<.001). There was no statistically significant

difference in the odds of having cognitive frailty between participants with short and modest sleep durations in the initial model adjusted for age, sex, and education. Compared with those who had a good sleep quality, participants who had a poor sleep quality were at a higher risk for cognitive frailty in both initial and the augmented models with further adjustments (initial model: OR 1.63, 95% CI 1.40 - 1.89; *P*<.001; augmented model: OR 1.41, 95% CI 1.21 - 1.64; *P*<.001; Table 2). In addition, RCS analysis revealed a J-shaped association between sleep duration and cognitive frailty at baseline (*P* for nonlinear<.001; Figure 2).

The results of sensitivity analyses are summarized in Multimedia Appendices 3 and 4. The associations between long sleep duration and baseline cognitive frailty remained consistent when adjusting for each chronic disease individually and when accounting for all chronic diseases simultaneously (Multimedia Appendix 3). For individuals younger than 80 years old, the association of long sleep duration with cognitive frailty became not statistically significant, whereas for individuals older than 80 years old, the association persisted. The association remained significant in sex-stratified models (Multimedia Appendix 4).

Table . Association between sleep duration and cognitive frailty at baseline.

	Model A[a]		Model B[b]		Model C[c]	
	OR[d] (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age	—[e]	—	1.12 (1.12 - 1.13)	<.001	1.12 (1.11 - 1.13)	<.001
Sex (female)	—	—	1.80 (1.54 - 2.11)	<.001	1.56 (1.30 - 1.87)	<.001
Not educated	—	—	1.78 (1.49 - 2.13)	<.001	1.57 (1.31 - 1.90)	<.001
Married and living with spouse	—	—	—	—	1.21 (0.97 - 1.52)	.09
Rural residence	—	—	—	—	0.86 (0.75 - 0.99)	.04
Economic dependence	—	—	—	—	2.24 (1.75 - 2.89)	<.001
Loneliness	—	—	—	—	1.82 (1.60 - 2.09)	<.001
Smoker	—	—	—	—	0.91 (0.76 - 1.10)	.33
Drinker	—	—	—	—	1.03 (0.87 - 1.22)	.73
Multimorbidity	—	—	—	—	1.47 (1.14 - 1.87)	.002
Poor sleep quality	1.62 (1.41 - 1.85)	<.001	1.63 (1.40 - 1.89)	<.001	1.41 (1.21 - 1.64)	<.001
Sleep duration (hours)						
Moderate (6 - 9)	reference	reference	reference	reference	reference	reference
Short (<6)	1.01 (0.82 - 1.23)	.95	1.04 (0.83 - 1.29)	.75	1.06 (0.85 - 1.32)	.61
Long (>9)	2.47 (2.17-2.82)	<.001	1.71 (1.48 - 1.97)	<.001	1.69 (1.46 - 1.95)	<.001

[a]Model was unadjusted.

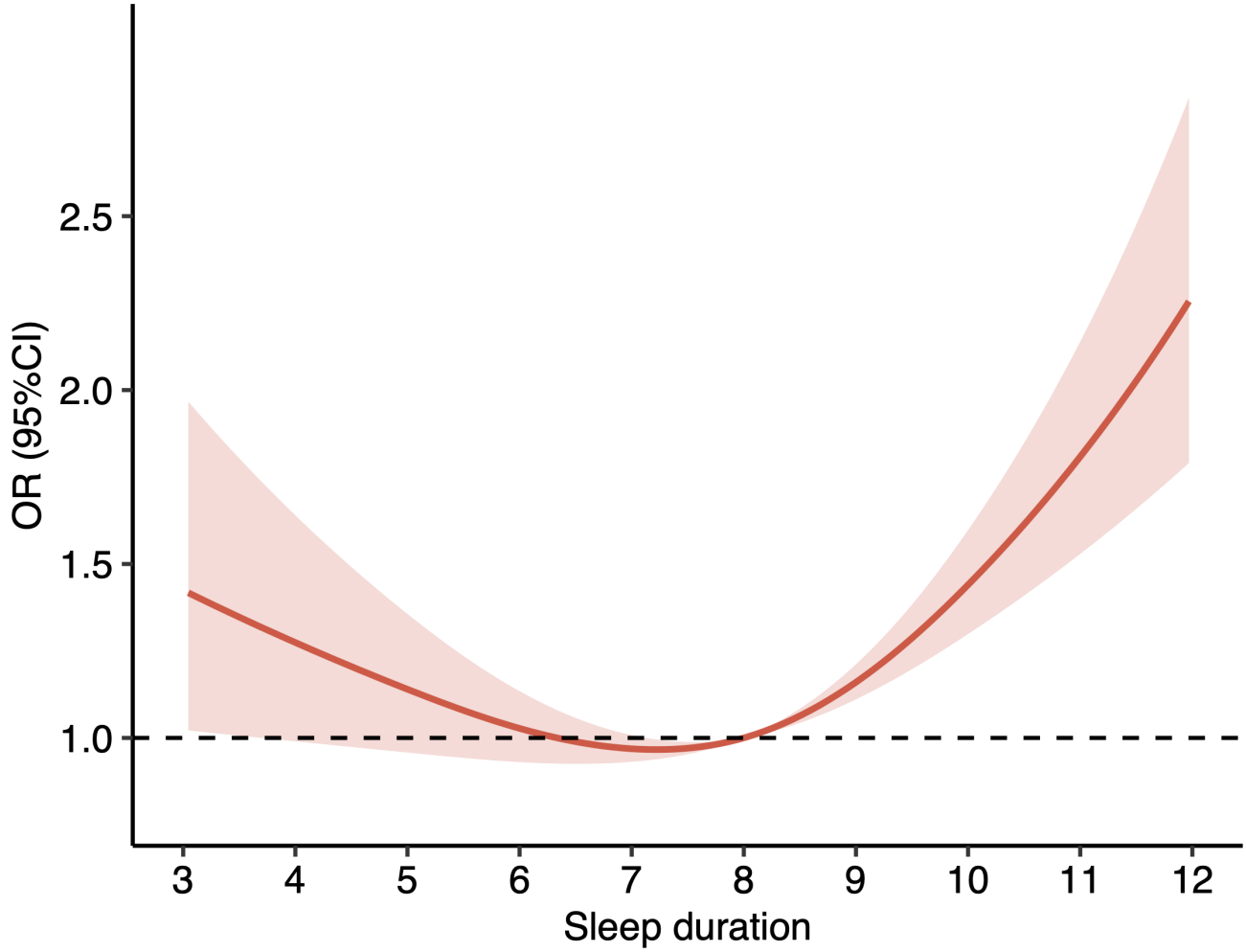
[b]Model was adjusted for age, sex, and education at baseline.

[c]Model was adjusted for age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity at baseline.

[d]OR: odds ratio.

[e]Not applicable.

Figure 2. Nonlinear association between sleep duration and cognitive frailty at baseline using restricted cubic spline. The logistic regression model was used to fit the data, with solid lines representing the estimated OR and shaded areas representing the corresponding 95% CIs. OR: odds ratio.



Association Between Sleep Duration and Incident Cognitive Frailty During Follow-Up

Over a mean of 6.7 (SD 2.6) years of follow up, 521 (10%) participants out of the 5201 participants developed cognitive frailty. In the initial model adjusted for age, sex, and education, as compared with participants who had moderate sleep duration at baseline, a higher risk of cognitive frailty was observed in participants who had long sleep duration at baseline (hazard ratio [HR] 1.32, 95% CI 1.07 - 1.62; *P*=.008; Table 3). The association persisted after further adjusting for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (HR 1.30, 95% CI 1.06 - 1.59; *P*=.01). There was no statistically significant difference in the hazard of incident cognitive frailty between participants with short sleep duration and moderate sleep duration in the initial model. Besides, there was no significant difference in the risk of incident cognitive frailty between participants with good and poor sleep quality. Additionally, RCS analysis showed no significant nonlinear association between sleep duration and cognitive frailty during follow up (*P* for nonlinear=.29; Figure 3).

Table . Association between sleep duration and incident cognitive frailty during follow up.

	Model A[a]		Model B[b]		Model C[c]	
	HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age	__[e]	—	1.09 (1.08 - 1.10)	<.001	1.09 (1.08 - 1.10)	<.001
Sex (female)	—	—	1.98 (1.60 - 2.45)	<.001	1.61 (1.25 - 2.08)	<.001
Not educated	—	—	1.84 (1.47 - 2.31)	<.001	1.72 (1.36 - 2.17)	<.001
Married and living with spouse	—	—	—	—	1.27 (1.01 - 1.59)	.04
Rural residence	—	—	—	—	1.03 (0.86 - 1.24)	.74
Economic dependence	—	—	—	—	1.47 (1.14 - 1.91)	.003
Loneliness	—	—	—	—	1.22 (1.01 - 1.47)	.04
Smoker	—	—	—	—	1.02 (0.80 - 1.30)	.85
Drinker	—	—	—	—	0.77 (0.61 - 0.97)	.03
Multimorbidity	—	—	—	—	1.33 (0.96 - 1.83)	.08
Poor sleep quality	1.32 (1.08 - 1.60)	.006	1.22 (1.00 - 1.49)	.05	1.14 (0.93 - 1.39)	.22
Sleep duration (hours)						
Moderate (6 - 9)	reference	reference	reference	reference	reference	reference
Short (<6)	0.96 (0.72 - 1.27)	.76	0.90 (0.68 - 1.21)	.49	0.94 (0.70 - 1.25)	.66
Long (>9)	1.74 (1.42 - 2.14)	<.001	1.32 (1.07 - 1.62)	.008	1.30 (1.06 - 1.59)	.01

[a]Model was unadjusted.

[b]Model was adjusted for age, sex, and education at baseline.

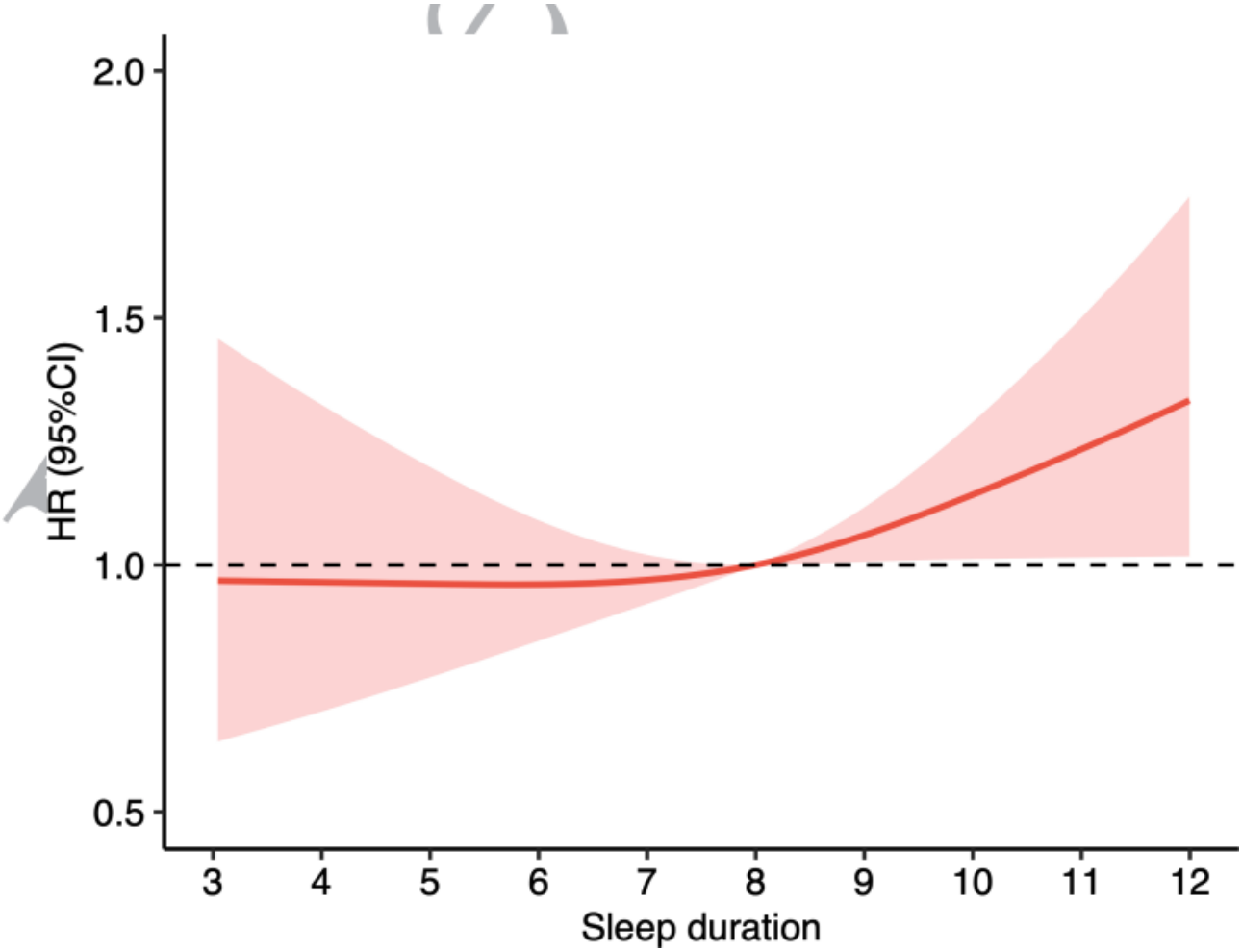
[c]Model was adjusted for age, sex, education, marital status, residence, economic status, loneliness, smoking status, drinking status and multimorbidity at baseline.

[d]HR: hazard ratio.

[e]Not applicable.

The results of sensitivity analyses are summarized in Multimedia Appendices 5 and 6. The associations between long sleep duration and baseline cognitive frailty status were consistent when adjusting for each chronic disease individually, or when adjusting for all chronic diseases simultaneously (Multimedia Appendix 5). For individuals younger than 80 years old, the association of long sleep duration with cognitive frailty became not statistically significant, whereas for individuals older than 80 years old, the association persisted. The association was not significant in male participants but remained statistically significant in female participants (Multimedia Appendix 6).

Figure 3. Nonlinear association between sleep duration and cognitive frailty during follow up using restricted cubic spline. The Cox proportional model was used to fit the data, with solid lines representing the estimated HR and shaded areas representing the corresponding 95% CIs. HR: hazard ratio.



Rate of Change in Sleep Duration and Cognitive Frailty

The mean annual rate of change in sleep duration was –0.02 (SD 0.87) hours/year. In the model adjusted for age, sex, education, sleep quality, and sleep duration at baseline, a faster increase in sleep duration per year was associated with a higher risk for cognitive frailty: for 1 hour/year increase in sleep duration, the HR was 1.14 (95% CI 1.03 - 1.27; *P*=.01). The association remained significant after further adjusting for marital status, residence, economic status, loneliness, smoking status, drinking status, and multimorbidity (HR 1.15, 95% CI 1.03 - 1.27; *P*=.01; Multimedia Appendix 7).

Discussion

Principal Findings

In this large community-based prospective study of over 10,000 Chinese older adults, we found that long sleep duration was consistently associated with cognitive frailty status at baseline, as well as an increased risk of cognitive frailty incident during follow up. We also found that a faster annual increase in sleep duration was associated with a higher risk of cognitive frailty during follow up.

The associations of sleep duration with physical frailty or cognitive decline are complex with mixed findings in prior research. A study using data from the National Health and Nutrition Examination Survey demonstrated that prolonged sleep duration (ie, ≥10 h) was associated with physical frailty in older adults, while a prospective study of 309 older Mexican adults showed that participants with either short sleep duration (ie, ≤5 h) or long sleep duration (ie, ≥9 h) had a higher risk of physical frailty [22,23]. Likewise, a Japanese cohort study of 623 older adults reported that long sleepers (>8 h) had a higher risk of cognitive impairment, while a cohort study found that self-reported short sleepers (<7 h) had an increased risk of cognitive decline [24,25]. However, few studies explored the association between sleep duration and cognitive frailty, that is, the coexistence of physical frailty and cognitive impairment. Two recent cross-sectional studies reported an association of long sleep duration with cognitive frailty among community-dwelling older adults and in older adults with heart failure [11,12]. In this study, we confirmed such a cross-sectional association between long sleep duration and cognitive frailty. More intriguingly, we showed consistent evidence from longitudinal analyses that participants with long sleep duration at baseline had an increased risk of developing cognitive frailty in the future. Of note, we did not find a significant association between short sleep duration and cognitive frailty in this study.

We also used the annual rate of change in sleep duration to assess the relationship between sleep duration changes and cognitive frailty. The results showed that a faster annual increase in sleep duration was associated with a higher risk of cognitive frailty, independent of the baseline sleep duration category. Notably, this study accounted for variations in follow-up years by using the annual rate of sleep duration change instead of the absolute change in sleep duration, minimizing classification bias and providing a more accurate representation of the relationship between sleep duration trajectories and cognitive frailty.

The mechanisms underlying the association of sleep duration with incident cognitive frailty remain unclear. However, several potential pathways can be suggested. First, inappropriate sleep duration (either shorter or longer than ideal) has been linked to cardiovascular health [26], which has been reported as a risk factor for both physical frailty and cognitive decline [27,28]. Second, increased sleep duration has also been linked with elevations in C-reactive protein, a marker for systemic inflammation, and interleukin-6, a proinflammatory biomarker [29,30]. Prior studies have shown that even low-level inflammation may pose an increased risk for the development of physical frailty and cognitive impairment [31,32]. Finally, long sleep duration may also be an indicator of or linked to disturbances in circadian control, the internal biological clock that prepares bodily responses to environmental light-dark changes. Circadian disturbances have been previously reported to be associated with both physical frailty and cognitive impairment or decline [33-36].

Several previous studies showed that sleep quality is closely related to impairments in both physical and cognitive domains [37,38]. In addition, 2 recent cross-sectional studies reported that poor sleep quality was associated with cognitive frailty in older Chinese adults living in nursing homes, and among Thai community-dwelling older adults during the COVID-19 restrictions [39,40]. Our observed cross-sectional association between sleep quality and cognitive frailty aligns with these previous findings. However, we did not find evidence of an association between sleep quality and the development of cognitive frailty in this study. It is important to note that the assessment of sleep quality in this study relied on a single question, which may be overly simplistic and may not fully capture the complexity of sleep behaviors. Future studies should use more comprehensive approaches to assess sleep health, considering its multidimensional nature, which includes factors such as duration, timing, efficiency, satisfaction, and alertness [41].

Strengths and Limitations

To the best of our knowledge, this study was the first to assess the association between sleep duration and incident cognitive frailty in a nationally representative sample of older Chinese adults. Strengths of this study include its longitudinal design and a large sample size. This study also has several limitations. First, the sleep duration and sleep quality were from self-reports and may be subjected to recall biases, especially in older adults and individuals with cognitive impairment. Future studies, ideally with an objective sleep monitoring approach (such as combining actigraphy and sleep diary or polysomnography), are warranted. Second, a growing body of studies have linked daytime napping behaviors with physical frailty, cognitive decline, Alzheimer dementia, or cognitive frailty [40,42,43]. Similarly, sleep disorders, such as sleep apnea, have been consistently linked with physical frailty, cognitive impairment, and dementia [10,44]. However, these potential sleep-related influences were not collected in the CLHLS. Future studies may consider incorporating these multidimensional structs when assessing sleep health and its relationships with physical and cognitive outcomes. Finally, the participants in this study were, on average, older than 80 years at baseline. In subgroup analyses, the association between long sleep duration and cognitive frailty was not statistically significant in participants younger than 80 years. This may be due to the limited number of cognitive frailty events in this subgroup, which could have resulted in insufficient statistical power to detect a significant association. Therefore, caution should be taken when translating our findings into younger populations.

Conclusions

In this large cohort study of over 10,000 older Chinese adults, long sleep duration was consistently associated with an increased risk of cognitive frailty, both cross-sectionally and longitudinally. These findings underscore the importance of sleep monitoring in aging populations. Future research using objective sleep assessments and investigating the underlying biological mechanisms is warranted to inform more effective prevention strategies.

Acknowledgments

This work is supported by the Key Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2022SJZD141 to JC) and the BrightFocus Foundation Alzheimer's Research Program (A2020886S to PL). PL is additionally supported by the Funds to Sustain Research Excellence from the Brigham Research Institute and a start-up fund from the Department of Anesthesia, Critical Care and Pain Medicine at Massachusetts General Hospital. CG is additionally supported by Alzheimer's Association (AARFD-22 - 928372). KH is additionally supported by the National Institute on Aging (RF1AG064312; R01AG083799) and a start-up fund from the Department of Anesthesia, Critical Care and Pain Medicine at Massachusetts General Hospital. LG is additionally supported by the Alzheimer's Association (AACSF-23 - 1148490).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics of participants by sleep duration groups
[DOCX File, 20 KB - aging_v8i1e65183_app1.docx]

Multimedia Appendix 2
Baseline characteristics of participants by incident cognitive frailty status during follow-up.
[DOCX File, 20 KB - aging_v8i1e65183_app2.docx]

Multimedia Appendix 3
Association between sleep and cognitive frailty at baseline with individual chronic disease adjustments.
[DOCX File, 19 KB - aging_v8i1e65183_app3.docx]

Multimedia Appendix 4
Association between sleep and baseline cognitive frailty stratified by age and sex.
[DOCX File, 19 KB - aging_v8i1e65183_app4.docx]

Multimedia Appendix 5
Association between sleep and incident cognitive frailty with individual chronic disease adjustments.
[DOCX File, 19 KB - aging_v8i1e65183_app5.docx]

Multimedia Appendix 6
Association between sleep and incident cognitive frailty stratified by age and sex.
[DOCX File, 19 KB - aging_v8i1e65183_app6.docx]

Multimedia Appendix 7
Rate of change in sleep duration and cognitive frailty.
[DOCX File, 18 KB - aging_v8i1e65183_app7.docx]

References

1. Bae JB, Han JW, Kwak KP, et al. Impact of mild cognitive impairment on mortality and cause of death in the elderly. J Alzheimers Dis 2018;64(2):607-616. [doi: 10.3233/JAD-171182] [Medline: 29914024]
2. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
3. Boyle PA, Buchman AS, Wilson RS, Leurgans SE, Bennett DA. Physical frailty is associated with incident mild cognitive impairment in community-based older persons. J Am Geriatr Soc 2010 Feb;58(2):248-255. [doi: 10.1111/j.1532-5415.2009.02671.x] [Medline: 20070417]
4. Song X, Mitnitski A, Rockwood K. Nontraditional risk factors combine to predict Alzheimer disease and dementia. Neurology (ECronicon) 2011 Jul 19;77(3):227-234. [doi: 10.1212/WNL.0b013e318225c6bc] [Medline: 21753161]
5. Kelaiditi E, Cesari M, Canevelli M, et al. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) International Consensus Group. J Nutr Health Aging 2013 Nov;17(9):726-734. [doi: 10.1007/s12603-013-0367-2]
6. Feng L, Nyunt MSZ, Gao Q, et al. Physical frailty, cognitive impairment, and the risk of neurocognitive disorder in the Singapore longitudinal ageing studies. GERONA 2016 Mar 24:glw050. [doi: 10.1093/gerona/glw050]
7. Solfrizzi V, Scafato E, Lozupone M, et al. Additive role of a potentially reversible cognitive frailty model and inflammatory state on the risk of disability: the Italian longitudinal study on aging. Am J Geriatr Psychiatry 2017 Nov;25(11):1236-1248. [doi: 10.1016/j.jagp.2017.05.018] [Medline: 28689645]
8. Li J, Vitiello MV, Gooneratne NS. Sleep in normal aging. Sleep Med Clin 2018 Mar;13(1):1-11. [doi: 10.1016/j.jsmc.2017.09.001] [Medline: 29412976]
9. Bubu OM, Brannick M, Mortimer J, et al. Sleep, cognitive impairment, and Alzheimer's disease: a systematic review and meta-analysis. Sleep 2017 Jan 1;40(1). [doi: 10.1093/sleep/zsw032] [Medline: 28364458]
10. Sun R, Xie Y, Jiang W, Wang E, Li X. Effects of different sleep disorders on frailty in the elderly: a systematic review and meta-analysis of observational studies. Sleep Breath 2023 Mar;27(1):91-101. [doi: 10.1007/s11325-022-02610-5] [Medline: 35416613]
11. Zhao Y, Lu Y, Zhao W, et al. Long sleep duration is associated with cognitive frailty among older community-dwelling adults: results from West China Health and Aging Trend study. BMC Geriatr 2021 Oct 27;21(1):608. [doi: 10.1186/s12877-021-02455-9] [Medline: 34706663]
12. Seo EJ, Son YJ. The prevalence of cognitive frailty and its association with sleep duration and depression among older adults with heart failure. Clin Gerontol 2024;47(3):416-425. [doi: 10.1080/07317115.2022.2125858] [Medline: 36117469]

13. Salinas-Rodríguez A, Manrique-Espinoza B, Moreno-Tamayo K, Guerrero-Zúñiga S. Trajectories of sleep duration and quality and their association with mild cognitive impairment, frailty, and all-cause mortality. Sleep Health 2024 Apr;10(2):240-248. [doi: 10.1016/j.sleh.2023.12.002] [Medline: 38238122]
14. Zeng Y. Chinese longitudinal healthy longevity survey and some research findings. Geriatrics Gerontol Int 2004 Sep;4(s1). [doi: 10.1111/j.1447-0594.2004.00147.x]
15. Zeng Y, Feng Q, Hesketh T, Christensen K, Vaupel JW. Survival, disabilities in activities of daily living, and physical and cognitive functioning among the oldest-old in China: a cohort study. Lancet 2017 Apr;389(10079):1619-1629. [doi: 10.1016/S0140-6736(17)30548-2]
16. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
17. Wang HY, Lv X, Du J, Kong G, Zhang L. Age- and gender-specific prevalence of frailty and its outcomes in the longevous population: the Chinese longitudinal healthy longevity study. Front Med (Lausanne) 2021;8:719806. [doi: 10.3389/fmed.2021.719806] [Medline: 34409056]
18. Yi Z, Vaupel JW. Functional capacity and self–evaluation of health and life of oldest old in China. J Soc Issues 2002 Jan;58(4):733-748. [doi: 10.1111/1540-4560.00287]
19. Li J, Cao D, Huang Y, et al. Sleep duration and health outcomes: an umbrella review. Sleep Breath 2022 Sep;26(3):1479-1501. [doi: 10.1007/s11325-021-02458-1] [Medline: 34435311]
20. Vatanabe IP, Pedroso RV, Teles RHG, et al. A systematic review and meta-analysis on cognitive frailty in community-dwelling older adults: risk and associated factors. Aging Ment Health 2022 Mar 4;26(3):464-476. [doi: 10.1080/13607863.2021.1884844]
21. Yuan C, Zhang Y. Modifiable risk factors of cognitive frailty in community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Knowl 2024 Apr;35(2):177-185. [doi: 10.1111/2047-3095.12437] [Medline: 37378541]
22. Baniak LM, Yang K, Choi J, Chasens ER. Long sleep duration is associated with increased frailty risk in older community-dwelling adults. J Aging Health 2020;32(1):42-51. [doi: 10.1177/0898264318803470] [Medline: 30270714]
23. Moreno-Tamayo K, Manrique-Espinoza B, Morales-Carmona E, Salinas-Rodríguez A. Sleep duration and incident frailty: the rural frailty study. BMC Geriatr 2021 Jun 16;21(1):368. [doi: 10.1186/s12877-021-02272-0] [Medline: 34134643]
24. Keil SA, Schindler AG, Wang MX, et al. Longitudinal sleep patterns and cognitive impairment in older adults. JAMA Netw Open 2023 Dec 1;6(12):e2346006. [doi: 10.1001/jamanetworkopen.2023.46006] [Medline: 38048131]
25. Kinoshita K, Otsuka R, Takada M, et al. Dietary amino acid intake and sleep duration are additively involved in future cognitive decline in Japanese adults aged 60 years or over: a community-based longitudinal study. BMC Geriatr 2023 Oct 11;23(1):653. [doi: 10.1186/s12877-023-04359-2] [Medline: 37821805]
26. Lloyd-Jones DM, Allen NB, Anderson CAM, et al. Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. Circulation 2022 Aug 2;146(5):e18-e43. [doi: 10.1161/CIR.0000000000001078] [Medline: 35766027]
27. Palta P, Griswold M, Ranadive R, et al. Midlife cardiovascular health and robust versus frail late-life status: the atherosclerosis risk in communities study. J Gerontol A Biol Sci Med Sci 2022 Jun 1;77(6):1222-1229. [doi: 10.1093/gerona/glab310] [Medline: 34661638]
28. Cho S, Yang PS, Kim D, et al. Association of cardiovascular health with the risk of dementia in older adults. Sci Rep 2022 Sep 19;12(1):15673. [doi: 10.1038/s41598-022-20072-3]
29. Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. Biol Psychiatry 2016 Jul 1;80(1):40-52. [doi: 10.1016/j.biopsych.2015.05.014] [Medline: 26140821]
30. Patel SR, Zhu X, Storfer-Isser A, et al. Sleep duration and biomarkers of inflammation. Sleep 2009 Feb;32(2):200-204. [doi: 10.1093/sleep/32.2.200]
31. Singh-Manoux A, Dugravot A, Brunner E, et al. Interleukin-6 and C-reactive protein as predictors of cognitive decline in late midlife. Neurology (ECronicon) 2014 Aug 5;83(6):486-493. [doi: 10.1212/WNL.0000000000000665] [Medline: 24991031]
32. Wang J, Maxwell CA, Yu F. Biological processes and biomarkers related to frailty in older adults: a state-of-the-science literature review. Biol Res Nurs 2019 Jan;21(1):80-106. [doi: 10.1177/1099800418798047] [Medline: 30198309]
33. Cai R, Gao L, Gao C, et al. Circadian disturbances and frailty risk in older adults. Nat Commun 2023 Nov 16;14(1):7219. [doi: 10.1038/s41467-023-42727-z] [Medline: 37973796]
34. Gao L, Li P, Gaykova N, et al. Circadian rest-activity rhythms, delirium risk, and progression to dementia. Ann Neurol 2023 Jun;93(6):1145-1157. [doi: 10.1002/ana.26617] [Medline: 36808743]
35. Li P, Gao L, Gaba A, et al. Circadian disturbances in Alzheimer's disease progression: a prospective observational cohort study of community-based older adults. Lancet Healthy Longevity 2020 Dec;1(3):e96-e105. [doi: 10.1016/s2666-7568(20)30015-5] [Medline: 34179863]
36. Li P, Hu K. Circadian disturbances and age-related disorders. Adv Biol (Weinh) 2023 Nov;7(11):e2300405. [doi: 10.1002/adbi.202300405] [Medline: 37986670]

37. Gildner TE, Liebert MA, Kowal P, Chatterji S, Snodgrass JJ. Associations between sleep duration, sleep quality, and cognitive test performance among older adults from six middle income countries: results from the Study on Global Ageing and Adult Health (SAGE). J Clin Sleep Med 2014 Jun 15;10(6):613-621. [doi: 10.5664/jcsm.3782] [Medline: 24932140]
38. Wai JT, Yu DF. The relationship between sleep-wake disturbances and frailty among older adults: a systematic review. J Adv Nurs 2020 Jan;76(1):96-108. [doi: 10.1111/jan.14231] [Medline: 31588595]
39. Griffiths J, Seesen M, Sirikul W, Siviroj P. Malnutrition, depression, poor sleep quality, and difficulty falling asleep at night are associated with a higher risk of cognitive frailty in older adults during the COVID-19 restrictions. Nutrients 2023 Jun 23;15(13):2849. [doi: 10.3390/nu15132849] [Medline: 37447178]
40. Liu S, Hu Z, Guo Y, Zhou F, Li S, Xu H. Association of sleep quality and nap duration with cognitive frailty among older adults living in nursing homes. Front Public Health 2022 Aug 25;10:963105. [doi: 10.3389/fpubh.2022.963105]
41. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298]
42. Li P, Gao L, Yu L, et al. Daytime napping and Alzheimer's dementia: a potential bidirectional relationship. Alzheimers Dement 2023 Jan;19(1):158-168. [doi: 10.1002/alz.12636] [Medline: 35297533]
43. Zhang Y, Zhou L, Ge M, Lin X, Dong B. Association between daytime nap duration and risks of frailty: findings from the China Health and Retirement Longitudinal Study. Front Public Health 2022;10:1098609. [doi: 10.3389/fpubh.2022.1098609] [Medline: 36777767]
44. Bubu OM, Andrade AG, Umasabor-Bubu OQ, et al. Obstructive sleep apnea, cognition and Alzheimer's disease: a systematic review integrating three decades of multidisciplinary research. Sleep Med Rev 2020 Apr;50:101250. [doi: 10.1016/j.smrv.2019.101250] [Medline: 31881487]

Abbreviations

C-MMSE: Chinese version of the Mini-Mental State Examination
CLHLS: Chinese Longitudinal Healthy Longevity Survey
HR: hazard ratio
OR: odds ratio
PH: proportional hazard
RCS: restricted cubic spline

Edited by F Puga; submitted 13.08.24; peer-reviewed by IC Wu, WQ Song; revised version received 10.03.25; accepted 20.03.25; published 23.04.25.

Please cite as:
Cai R, Chao J, Gao C, Gao L, Hu K, Li P
Association Between Sleep Duration and Cognitive Frailty in Older Chinese Adults: Prospective Cohort Study
JMIR Aging 2025;8:e65183
URL: https://aging.jmir.org/2025/1/e65183
doi:10.2196/65183

Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study

Natthanaphop Isaradech[1,2*], MD; Wachiranun Sirikul[1,3,4*], MD, MSc; Nida Buawangpong[5], MD; Penprapa Siviroj[1], DrPH; Amornphat Kitro[1,2], MD, MCTM

[1]Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

[2]Environmental and Occupational Medicine Excellence Center, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[3]Center of Data Analytics and Knowledge Synthesis for Health Care, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[4]Department of Biomedical Informatics and Clinical Epidemiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

[*]these authors contributed equally

Corresponding Author:
Wachiranun Sirikul, MD, MSc
Department of Community Medicine, Faculty of Medicine, Chiang Mai University, 110, Intrawarorot Road, Meaung, Thailand

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2025/1/e75690

Abstract

Background: Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's physical function resulting in increased morbidity and mortality when exposed to acute stressors. Early identification and management can reverse individuals with frailty to being robust once more. However, we found no integration of machine learning (ML) tools and frailty screening and surveillance studies in Thailand despite the abundance of evidence of frailty assessment using ML globally and in Asia.

Objective: We propose an approach for early diagnosis of frailty in community-dwelling older individuals in Thailand using an ML model generated from individual characteristics and anthropometric data.

Methods: Datasets including 2692 community-dwelling Thai older adults in Lampang from 2016 and 2017 were used for model development and internal validation. The derived models were externally validated with a dataset of community-dwelling older adults in Chiang Mai from 2021. The ML algorithms implemented in this study include the k-nearest neighbors algorithm, random forest ML algorithms, multilayer perceptron artificial neural network, logistic regression models, gradient boosting classifier, and linear support vector machine classifier.

Results: Logistic regression showed the best overall discrimination performance with a mean area under the receiver operating characteristic curve of 0.81 (95% CI 0.75 - 0.86) in the internal validation dataset and 0.75 (95% CI 0.71 - 0.78) in the external validation dataset. The model was also well-calibrated to the expected probability of the external validation dataset.

Conclusions: Our findings showed that our models have the potential to be utilized as a screening tool using simple, accessible demographic and explainable clinical variables in Thai community-dwelling older persons to identify individuals with frailty who require early intervention to become physically robust.

(JMIR Aging 2025;8:e62942) doi:10.2196/62942

KEYWORDS

aged care; gerontology; geriatric; old; aging; clinical decision support; delivering health information and knowledge to the public; diagnostic systems; digital health; epidemiology; surveillance; diagnosis; frailty; machine learning; prediction; predictive; AI; artificial intelligence; Thailand; community dwelling; health care intervention; patient care

Introduction

The world population is moving toward an aging society. As health care technology improves, people are expected to live longer and healthier [1]. According to the World Health Organization, the population aged ≥60 years will increase from 1 billion in 2020 to 2.1 billion in 2050 and the number of people aged ≥80 years will reach 426 million in 2050 [2]. Researchers predicted that the proportion of people in Thailand aged ≥60

years would be more than 20% of the population in 2025 and more than 30% in 2031 [3,4].

The prevalence of frailty is high among older adults aged ≥60 years [5]. Global frailty prevalence ranges from approximately 10% to 12% [6-11]. The percentage varies by age, gender, and frailty classification tool. In Thailand, frailty prevalence was 22.1%, which is twice the global frailty prevalence, according to the Thai National Health Examination Survey cohort in 2018. Specifically, Thailand's northern region frailty prevalence was found to be 15% [12,13]. This creates concerns about the increasing aging population in Thailand.

Frailty is defined as a clinical state of increased vulnerability due to the age-associated decline of an individual's body resulting in increased morbidity and mortality when exposed to everyday or acute stressors [14,15]. This clinical syndrome is associated with decreased quality of life [7], slow gait speed [16], more depressive symptoms, higher BMI, reduced cognitive function [17], decreased strength [12], and increased risk of fall, hospital re-admission, and all-cause mortality in the older adult population [13,18-20]. Frailty has become a crucial research topic because this clinical syndrome can be reversed. Studies have shown that early detection and intervention can revert individuals from a frail to a fit state [21-23].

However, incorporating frailty evaluation into clinical practice in a primary care context is challenging due to increased administrative tasks, time limitations, and a lack of diagnostic effectiveness [15,24,25]. In the age of technology, health informatics has become an important role in health care research [26,27]. Several studies have applied information technology to frailty detection in primary care settings using machine learning models and artificial intelligence [28-30]. A study from Canada showcased an efficient frailty identification tool using the XGBoost machine learning model. Features used for the model were medication, medical billing codes, and other primary care clinical data [31]. Another example showed the development of a predictive machine learning model for frailty conditions based on a database of demographic data and clinical characteristics [32]. In China, a study simplified the Frailty Index assessment for older individuals using machine learning techniques and showed that logistic regression was the best performing and most interpretable model, with a mean area under the receiver operating characteristic curve (AUC) of 0.974 in the internal validation dataset [33]. Another study in China also developed and validated models using data from 6997 older adult participants to predict frailty risk, with random forest (RF) and logistic regression (LR) achieving AUC values of 0.77 and 0.76, respectively [34].

Despite the abundance of evidence of conducting frailty assessments using machine learning globally and in Asia, we have found no integration of data science tools and frailty screening and surveillance studies in Thailand. Most studies focused on the risk factors and their association with frailty syndrome but failed to show application in real-world settings [12,13,35,36]. This leads to our research question, "Can machine learning models predict frailty in community-dwelling older adults in Northern Thailand?" Therefore, we propose an approach to frailty detection using a machine learning model in the community-dwelling population from Lampang and Chiang Mai, Thailand, to effectively screen frailty status among Thai community-dwelling older adults and to help decrease clinicians' burden of work.

Methods

Source of Data

Development and Internal Validation Datasets

The datasets were derived from a cross-sectional study carried out in Lampang in 2016 and 2017; it is a northern Thai province with one of the highest aging indexes [37]. This study included older adults aged ≥60 years. Those with dementia (as determined by the Thai Mental State Examination), blindness, deafness, bedridden status, disabilities, or severe acute diseases were excluded.

To represent urban (8 villages), semiurban (8 villages), and rural (8 villages) communities, 24 villages in 3 districts were chosen. The records of the primary care unit were used to compile lists of community-dwelling older adults. The details of the data collection in this study are published elsewhere [36]. A total of 2228 older adults from this study were derived for model development and internal validation.

External Validation Dataset

The datasets used for external validation were derived from a cross-sectional study in Kuamung (suburban), Sankampang, Chiang Mai, Thailand, in 2021. The participants were included and excluded from the study with the same criteria as described in the internal validation datasets. A total of 464 older adults from this study were used for external validation of the derived models.

Predictors

Characteristics and Demographics

Assessments were conducted through a questionnaire at the participant's residence. The questionnaire included questions about sociodemographic information (age, household living arrangement, gender, and education level), self-reported medical diagnoses (such as hypertension, diabetes mellitus, and heart disease), level of physical activity per week, and exhaustion.

Anthropometric Variables

Anthropometric measures included BMI, waist circumference, and calf circumference (CC). A handheld dynamometer (Takei TKK5001) was used to assess handgrip strength. Height and calf circumference were measured with standard tape (Tajima brand, PIT-20BL model), and weight was measured with a calibrated weighing scale (Shaper Disney). The CC was measured over the unclothed area at the maximum diameter on the left leg. The tape was wrapped snugly around the calf and measured to the nearest 0.1 cm. All measurements were administered by 10 qualified field investigators, and the measurements were standardized by the principal investigator.

Outcome Variable

Frailty was evaluated based on Fried's phenotype [38], which includes the following five criteria:

1. Unintentional weight loss: The participant will be asked "In the last ten years, have you lost more than 10 pounds intentionally (not due to diet or exercise)?" If yes, then frail for weight loss criterion. At follow-up, weight loss was calculated as: (weight in previous year – current measured weight)/(weight in previous year)=K. If K≥0.05 and the subject did not report that he/she was trying to lose weight (ie, unintentional weight loss of at least 5% of previous year's body weight), then frail for weight loss=yes.
2. Exhaustion, assessed with self-report using Fried's method of assessment. The participant was first asked to self-assess whether she/he felt exhausted. If yes, she/he would be asked to rate the severity of the exhaustion. Ratings of 2-4 suggested a positive assessment.
3. Physical activity, based on the short version of the Minnesota Leisure time Activity questionnaire, which asked about walking, chores (moderately strenuous), mowing the lawn, raking, gardening, hiking, jogging, biking, exercise cycling, golf, single tennis, doubles tennis, racquetball, and calisthenics. For men, those with <383 kcals of physical activity per week were frail. For women, those with <270 kcals of physical activity per week were frail.
4. Walk time, stratified by gender and height (gender-specific cutoff for medium height).
   - Height ≤173 cm and walk time ≥7 seconds for men.
   - Height >173 cm and walk time ≥6 seconds for men.
   - Height ≤159 cm and walk time ≥7 seconds for women.
   - Height >159 cm and walk time ≥6 seconds for women.
5. Grip strength, stratified by gender and BMI quartiles.
   a. For men:
      - BMI ≤24 and grip strength ≤29 for men
      - BMI 24.1 - 26 and grip strength ≤30
      - BMI 26.1 - 28 and grip strength ≤30
      - BMI >28 and grip strength ≤32
   b. For women:
      - BMI ≤23 and grip strength ≤17
      - BMI 23.1-26 and grip strength ≤17.3
      - BMI 26.1-29 and grip strength ≤18
      - BMI <29 and grip strength ≤21

Older adults with 3 or more phenotypes were considered to have physical frailty, and those who had 1 or 2 phenotypes were classified as prefrailty. According to the study design of the derived datasets, the outcome was concurrently measured with the predictors.

Sample Size Calculation

The sample size for model development was calculated using the *pmsampsize* package via STATA (version 16; StataCorp). A minimum sample size required 16.41 events per predictor, which was 1519 patients with 263 frailty cases based on a maximum candidate predictor of 16, frailty prevalence of 17.29% in the dataset for model development, small overfitting defined as expected shrinkage of predictor effects by 10% or less, a small difference in the developed model's apparent and optimism-adjusted values of 0.15 ($R^2_{Nagelkerke}$), as suggested by Riley et al [39]. For the external validation, the sample size was calculated using the *pmvalsampsize* package via STATA. From the available data of 464 patients with 192 frailty cases, it was able to precisely estimate confidence interval widths of observed/expected statistic of 0.22, calibration slope of 0.60, and concordance statistic of 0.10 [40].

Statistical Analysis Methods and Synthesis of the Results

Data exploration was performed using descriptive statistics to determine the data quantity, quality, and distribution. A frequency and a percentage were used to describe categorical variables. For continuous variables, a mean with SD was used for parametrically distributed data, and a median with an IQR was used for nonparametrically distributed data. The comparison of characteristics between the frail and nonfrail populations was performed using the chi-square test for categorical variables, the independent *t* test for parametric variables, and the rank-sum test for nonparametric variables. The methods for feature selection, model development, and validation are described in the section "Model Development and Validation". Estimates of model discrimination and optimism are reported as the mean AUC with 95% CI across all repetitions of cross-validation for the internal validation and the AUC with 95% CI from 1000 bootstrapping samples for the external validation. To further explain model performance, we also created model calibration plots and calculated secondary metrics of prediction models, including the confusion matrix and specificity, sensitivity, and predictive values.

Missing Data and Imputation

There was only missing data in participants' age in the internal validation dataset (4/2228, 0.18%); therefore, a complete case analysis was performed on the dataset. The external validation dataset contained missing data, including age (2/464, 0.43%), BMI (11/464, 2.37%), waist and calf circumference (3/464, 0.64%), handgrip strength (2/464, 0.43%), and frailty status (4/464, 0.86%). We performed multiple imputations of missing data, except for frailty status (the outcome of interest), using the predictive mean matching imputation with 5 nearest neighbors via the *KNNImputer* from the *Scikit-Learn library 1.1.2*. Four participant records that did not have frailty status were removed (list-wise deletion).

Model Development and Validation

Feature Selection

The classification models were developed using the variables from the derived datasets. The variables for model development were selected using both a data-driven method and domain expertise. For a data-driven selection, a multivariable logistic regression with stepwise backward elimination was performed to determine statistically significant variables (*P* value less than .20). For the *P* value threshold of .20, we used a higher threshold to give priority to clinical reasoning in selecting variables by domain expertise and associated factors from the previous studies, along with statistical significance, which would allow more important variables to be entered into the model.

Rebalancing Data Strategy

We resolved unbalanced classification for the dataset's minority class by using the Synthetic Minority Oversampling Technique (SMOTE) to enhance model decision boundaries via the *imblearn.over_sampling.SMOTE* package [41,42].

Model Development and Internal Validation

The model development and validation were performed using Python (version 3.9; Python Software Foundation). The machine learning algorithms implemented in this study include the k-nearest neighbors (KNN) algorithm, RF machine learning algorithms, multilayer perceptron artificial neural network (MLP), gradient boosting classifier (GBC), linear support vector machine classifier (SVM), and LR models via the *Scikit-Learn library 1.1.2*. The hyperparameters were determined by using a grid search via the *GridSearch CV* package with 10-fold cross-validation on the derived dataset to determine the parameters of each model that led to the best discriminative performance. For 10-fold cross-validation, the derived dataset was divided into 10 folds of data and repeated 10 times to perform model training and testing. For each iteration, 9 folds of data were used to train the model and then it was tested with the remaining fold to ensure that almost all the derived data were used to train and test the models. The discriminative performance of the derived models was assessed by computing a confusion matrix and sensitivity, specificity, and predictive values, as well as AUC with a 95% CI.

The model calibration was evaluated using the calibration plot, which indicated the congruence between the observed proportion of the actual probability of outcome and the mean predicted probability (MPP) from the derived models.

External Validation

The derived models were validated again with the external validation dataset to determine the model optimism and calibration. The discriminative performance and the model optimism were re-evaluated and presented by discriminative performance matrices and a 95% CI of AUC from the 1000 bootstrapping samples, respectively. The model calibration using the external validation dataset was re-evaluated using the calibration plot.

Ethical Considerations

This study complies with the research with exemption category and has been certified by the Research Ethics Committee of the Faculty of Medicine, Chiang Mai University (study code: COM-2565 - 09159, number EXEMPTION 9159/2022). We requested a waiver of informed consent because this study is a retrospective analysis that exclusively utilizes anonymized secondary data from our research database, without collecting any additional information from medical records or other sources. All participants have previously provided informed consent for the primary data collection as described elsewhere [35,36,43]. All personal patient data were anonymized by removing citizen ID numbers, hospital numbers, addresses, and contact information from the dataset. The investigator cannot trace or identify individuals. The study results were reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence (TRIPOD+AI) statement [44].

Results

Baseline Characteristics of Participants in Each Dataset

Baseline characteristics of participants in this study are shown in Table 1. The participants in the internal validation dataset had a mean age of 71.0 years. Most of the participants were male and had finished primary school. Among 2228 old adults, 2160 lived with either their spouse, relative, or children while the others were living alone. The average BMI of the participants was 32.6 (SD 7.4) kg/m^2. The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 45.26%, 19.67%, 16.21%, and 4.31%, respectively. The averages of waist circumference, calf circumference, handgrip strength, and walk time were 83.40 (SD 11.23) cm, 32.40 (SD 4.35) cm, 32.40 (SD 6.68) kg, and 6.42 (SD 2.01) minutes, respectively. Overall, 9.79% of the participants were exhausted and 16.3% had physical activity higher than 150 minutes per week. The prevalence of frailty was 17.3% (n=385).

Table . The characteristics of participants in the development and internal validation datasets and the external validation dataset.

Characteristics	Development and internal validation datasets (Lampang, 2016 - 2017; N=2228)	External validation datasets (Chiang Mai, 2021; N=464)	*P* value
Age (years), mean (SD)	70.96 (7.49)	70.68 (5.58)	.45
Gender, n (%)			
Male	1569 (70.45)	193 (41.59)	<.001
Female	658 (29.55)	271 (58.41)	
Household living arrangement, n (%)			
Living alone	160 (7.18)	42 (9.05)	<.001
Living with spouse	1177 (52.85)	9 (1.94)	
Living with children	823 (36.96)	283 (60.99)	
Living with relatives or others	67 (3.01)	130 (28.02)	
BMI (kg/m^2), mean (SD)	32.64 (7.40)	22.73 (3.89)	.001
Education, n (%)			
No education	192 (8.62)	12 (2.61)	<.001
Primary school	1745 (78.36)	398 (86.52)	
Secondary school or higher	290 (13.02)	50 (10.87)	
Underlying diseases, n (%)			
Hypertension	1008 (45.26)	237 (51.08)	.02
Dyslipidemia	438 (19.67)	79 (17.03)	.19
Type 2 diabetes mellitus	361 (16.21)	78 (16.81)	.75
Heart diseases	96 (4.31)	18 (3.88)	.68
Anthropometric variables, mean (SD)			
Waist circumference (cm)	83.40 (11.23)	81.16 (10.80)	<.001
Calf circumference (cm)	32.40 (4.35)	32.85 (4.55)	.04
Walk time (min), mean (SD)	6.42 (2.01)	8.50 (5.38)	<.001
Exhaustion, n (%)	218 (9.79)	189 (40.74)	<.001
Adequate level of physical activity as defined by the World Health Organization, n (%)	363 (16.30)	104 (22.42)	<.001
Grip strength (kg), mean (SD)	32.40 (6.68)	19.87 (7.29)	<.001
Frailty, n (%)	385 (17.29)	192 (41.74)	<.001

In the external validation dataset, the participants had a mean age of 70.68 years. Most of the participants were male and had finished primary school. Among 464 persons, 422 lived with either their spouse, relative, or children, while the others were living alone. The average BMI of the participants was 22.73 (SD 3.89) kg/m^2. The prevalence of hypertension, dyslipidemia, type 2 diabetes mellitus, and heart disease were 51.08%, 17.03%, 16.81%, and 3.88%, respectively. The averages of waist circumference, calf circumference, grip strength, and walk time were 81.16 (SD 10.80) cm, 32.85 (SD 4.55) cm, 19.87 (SD 7.29) kg, and 8.50 (SD 5.38) minutes, respectively. The prevalence of participants with exhausted state was 40.74%, and 22.42% of the participants participated in physical activity more than 150 minutes per week.

Comparing the variables between the 2 datasets, we found that gender, household living arrangement, BMI, education, waist circumference, walk time, exhaustion, grip strength, and level of physical activity were significantly different.

Model Development

We used a dataset from Lampang (2016 - 2017; N=2228) for model development. The association between candidate predictors and frailty by univariate analysis is reported in Table 2. Feature selection for model development was selected by a backward elimination approach via a multivariable logistic regression and expert judgment. We chose the following features as model predictors: age, gender, status, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Finally, the derived data included 385 participants with frailty and 1842 participants

without frailty, and all candidate predictors, as shown in Table 2, were used in model development.

Table . The comparison of participants' characteristics and their associations with frailty in the development and internal validation datasets.

Characteristics	Adjusted odds ratio – full model (95% CI)	*P* value	Adjusted odds ratio – reduced model (95% CI)	*P* value
Age (years)	1.04 (1.02 - 1.07)	<.001	1.09 (1.06 - 1.11)	<.001
Gender				
Male	Reference		Reference	
Female	1.93 (1.26 - 2.94)	.002	0.90 (0.63-1.27)	.55
Status				
Living alone	Reference			
Living with others	0.92 (0.46 - 1.86)	.82	0.87 (0.45 - 1.68)	.68
BMI (kg/m^2)	0.92 (0.90 - 0.95)	<.001	0.91 (0.88 - 0.93)	<.001
Education				
No education	Reference			
Primary school	1.00 (0.59 - 1.70)	.99		
Secondary school or higher	1.34 (0.65 - 2.77)	.42		
Underlying diseases				
Hypertension	1.47 (1.03 - 2.10)	.03	1.27 (0.92 - 1.75)	.15
Dyslipidemia	1.22 (0.78 - 1.91)	.38	1.45 (0.97 - 2.17)	.07
Type 2 diabetes mellitus	0.84 (0.54 - 1.31)	.44		
Heart disease	1.03 (0.45 - 2.34)	.95		
Anthropometric variables				
Waist circumference (cm)	1.02 (1.00 - 1.04)	.046	1.02 (1.00 - 1.04)	.02
Calf circumference (cm)	1.05 (1.01 - 1.09)	.01	1.04 (1.01 - 1.08)	.02
Walk time (min)	1.89 (1.71 - 2.10)	<.001	1.95 (1.78 - 2.13)	<.001
Exhaustion	20.39 (12.61 - 32.94)	<.001	28.23 (18.22 - 43.73)	<.001
Adequate level of physical activity as defined by the World Health Organization	4.49 (3.06 - 6.59)	<.001		
Grip strength	0.83 (0.79 - 0.85)	<.001		

However, the model appeared to be poor in frailty prediction performance and imbalanced as the classifiers intended to classify only the majority class (accuracy paradox). Therefore, a rebalancing strategy by SMOTE was applied to counter this problem. The oversampling data were generated and rebalanced the minority group in a 1:1 ratio. The details of the final model's hyperparameters using *GridSearch CV* are presented in Table S1 in Multimedia Appendix 1.

Adjusted odds ratios (aORs) of the full model were obtained from a multivariable logistic regression with all features. aORs of the reduced model were obtained from a multivariable logistic regression with stepwise backward elimination (P<.10) and feature selection based on the domain expertise.

Discrimination Performance of Internal Validated Models

We evaluated the model's performance by 10-fold cross-validation. The discrimination performances of the models are presented in Table 3 and Figure 1. The overall discrimination performance was presented by mean AUC 10-fold cross-validation. The KNN model achieved the highest overall performance with a mean AUC of 0.85 (95% CI 0.82 - 0.88), followed by MLP (AUC 0.81, 95% CI 0.72 - 0.89), LR (AUC 0.81, 95% CI 0.75 - 0.86), and SVM (AUC 0.75, 95% CI 0.75 - 0.86), respectively. In addition, the KNN model had the highest sensitivity and specificity (89% and 76%, respectively). RF had the lowest discrimination performance with almost all metrics. Other metrics that were not affected by data rebalancing, like positive predictive value (PPV) and negative predictive value (NPV), were also used to express the performance of the models. The best performances for both

PPV and NPV were found in the KNN model (0.79 and 0.87, respectively).

Table . Discrimination and optimism of internal validated models with 95% CIs.

Models	AUC[a], mean (95% CI)	Predictive values		Sensitivity	Specificity
		Positive	Negative		
LR[b]	0.81 (0.75 - 0.86)	0.74 (0.68 - 0.80)	0.74 (0.69 - 0.78)	0.74 (0.69 - 0.79)	0.73 (0.65 - 0.81)
KNN[c]	0.85 (0.82 - 0.88)	0.79 (0.76 - 0.82)	0.88 (0.86 - 0.90)	0.89 (0.87 - 0.91)	0.76 (0.71 - 0.81)
RF[d]	0.70 (0.60 - 0.79)	0.65 (0.58 - 0.73)	0.70 (0.63 - 0.76)	0.79 (0.74 - 0.83)	0.55 (0.39 - 0.71)
MLP[e]	0.81 (0.72 - 0.89)	0.73 (0.64 - 0.81)	0.75 (0.68 - 0.81)	0.78 (0.71 - 0.84)	0.69 (0.57 - 0.81)
GBC[f]	0.74 (0.72 - 0.89)	0.66 (0.59 - 0.74)	0.70 (0.63 - 0.78)	0.78 (0.73 - 0.83)	0.58 (0.43 - 0.73)
SVM[g]	0.75 (0.75 - 0.86)	0.74 (0.68 - 0.80)	0.73 (0.69 - 0.78)	0.74 (0.69 - 0.79)	0.73 (0.65 - 0.81)

[a]AUC: area under the receiver operating characteristic curve.

[b]LR: logistic regression.

[c]KNN: k-nearest neighbors.

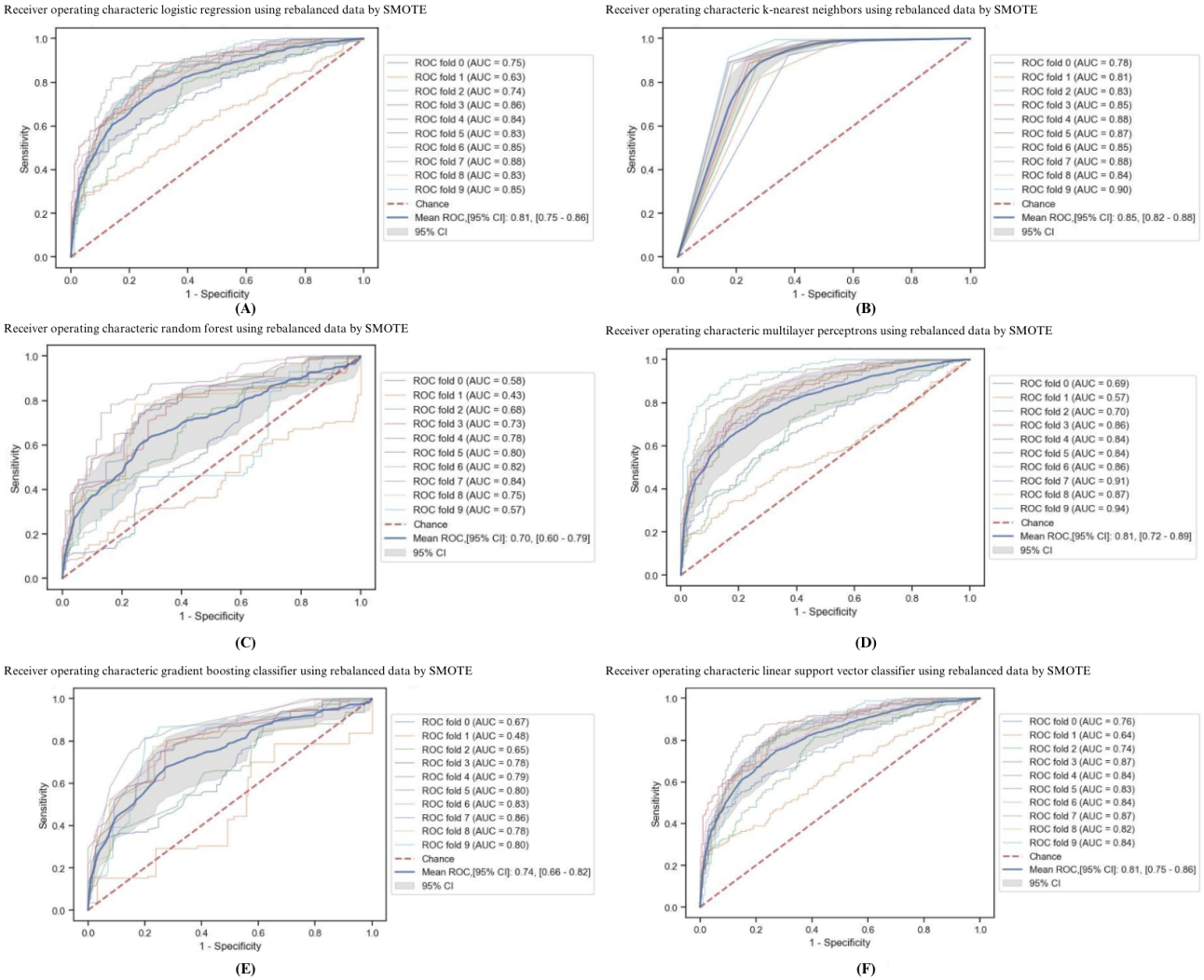
[d]RF: random forest.

[e]MLP: multilayer perceptron artificial neural network.

[f]GBC: gradient boosting classifier.

[g]SVM: linear support vector machine classifier.

Figure 1. Receiver operating characteristic curves from 10-fold cross-validation of the rebalanced learning classifiers by SMOTE: (A) logistic regression model: mean AUC 0.81 (95% CI 0.75 - 0.86); (B) k-nearest neighbors model: mean AUC 0.85 (95% CI 0.82 - 0.88); (C) random forest model: mean AUC 0.70 (95% CI 0.60 - 0.79); (D) multilayer perceptron model: mean AUC 0.81 (95% CI 0.72 - 0.89); (E) gradient boosting classifier model: mean AUC 0.74 (95% CI 0.66 - 0.82); (F) linear support vector machine classifier model: mean AUC 0.75 (95% CI 0.75 - 0.86). ROC: receiver operating characteristic curve; SMOTE: synthetic minority oversampling technique.



Discrimination Performance of External Validated Models

We validated our trained model with an external validation dataset to evaluate the bias and variance of our trained models. SMOTE was not applied to the dataset since it was already well-balanced between participants with and without frailty. The performance of machine learning models validated by the external validation dataset is shown in Table 4.

Table . Discrimination and optimism of external validated models.

Models and model prediction		True label (frailty/non-frailty)	AUC[a], mean (95% CI)	Predictive values		Sensitivity	Specificity
				Positive	Negative		
Logistic regression							
	Frailty	140/65	0.75 (0.71 - 0.78)	0.68	0.80	0.73	0.76
	Nonfrailty	52/207					
K-nearest neighbors							
	Frailty	41/35	0.54 (0.51 - 0.57)	0.54	0.61	0.21	0.87
	Nonfrailty	151/237					
Random forest							
	Frailty	137/63	0.75 (0.71 - 0.78)	0.69	0.79	0.71	0.77
	Nonfrailty	55/209					
Multilayer perceptron artificial neural network							
	Frailty	98/38	0.68 (0.65 - 0.72)	0.72	0.71	0.51	0.86
	Nonfrailty	94/234					
Gradient boosting classifier							
	Frailty	89/31	0.69 (0.65 - 0.72)	0.74	0.70	0.46	0.89
	Nonfrailty	103/241					
Linear support vector machine classifier							
	Frailty	131/60	0.73 (0.70 - 0.77)	0.69	0.78	0.68	0.78
	Nonfrailty	61/212					

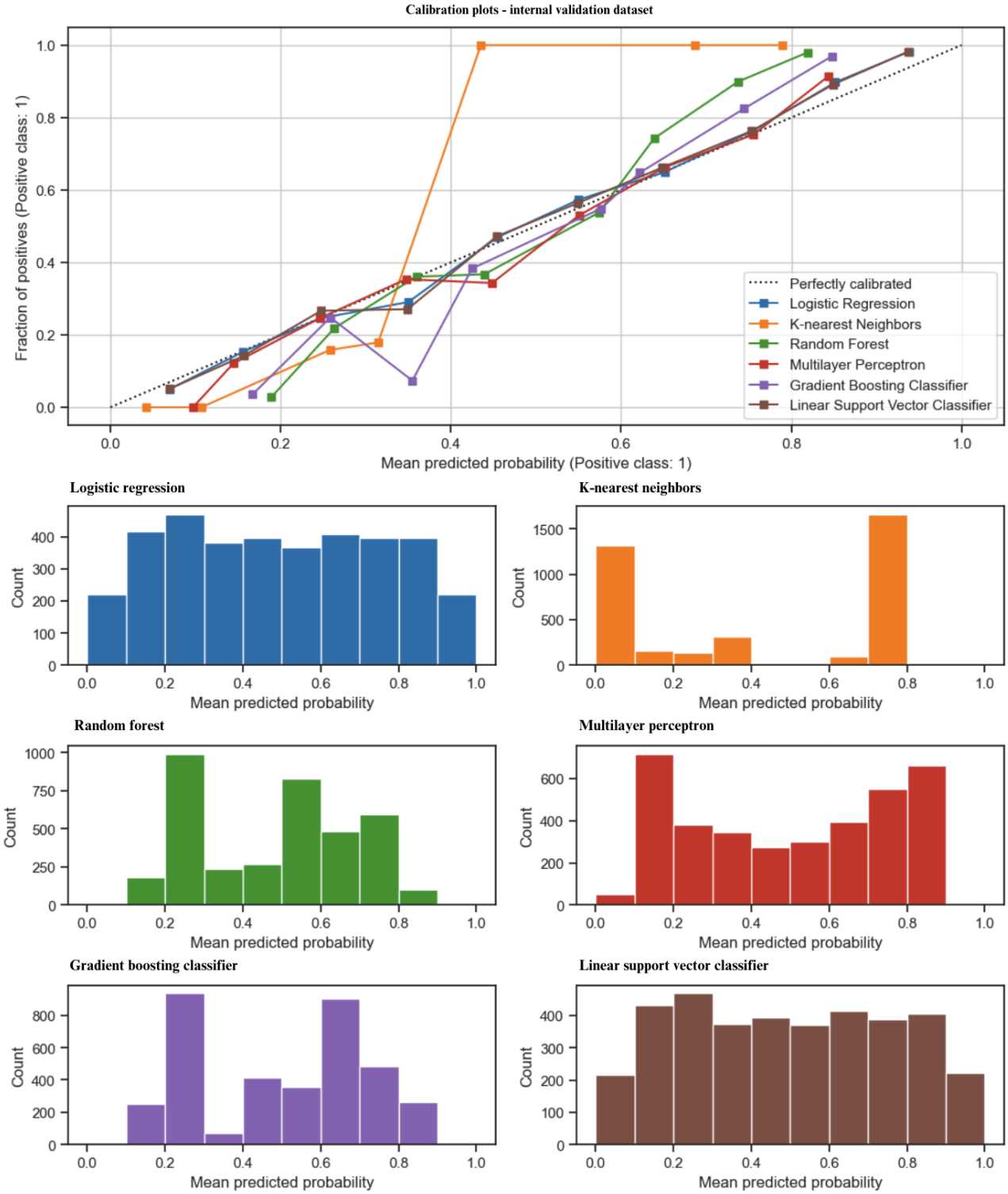
[a]AUC: area under the receiver operating characteristic curve.

The overall discrimination performance was presented by mean AUC and 95% CI via 1000-fold bootstrapping. The LR and RF model achieved the highest overall performance with mean AUC 0.75 (95% CI 0.71 - 0.78), followed by SVM (mean AUC 0.73, 95% CI 0.70 - 0.77), and GBC (mean AUC 0.69, 95% CI 0.65 - 0.72), respectively. The LR model had the highest sensitivity (73%) and the MLP model had the highest specificity (86%). KNN had the lowest discrimination performance with all metrics. Other metrics that were not affected by data rebalancing, like PPV and NPV, were also used to express the performance of the models. The best performance for PPV was found in the GBC model (0.74); for NPV, it was found in the LR model (0.80).

Model Calibration

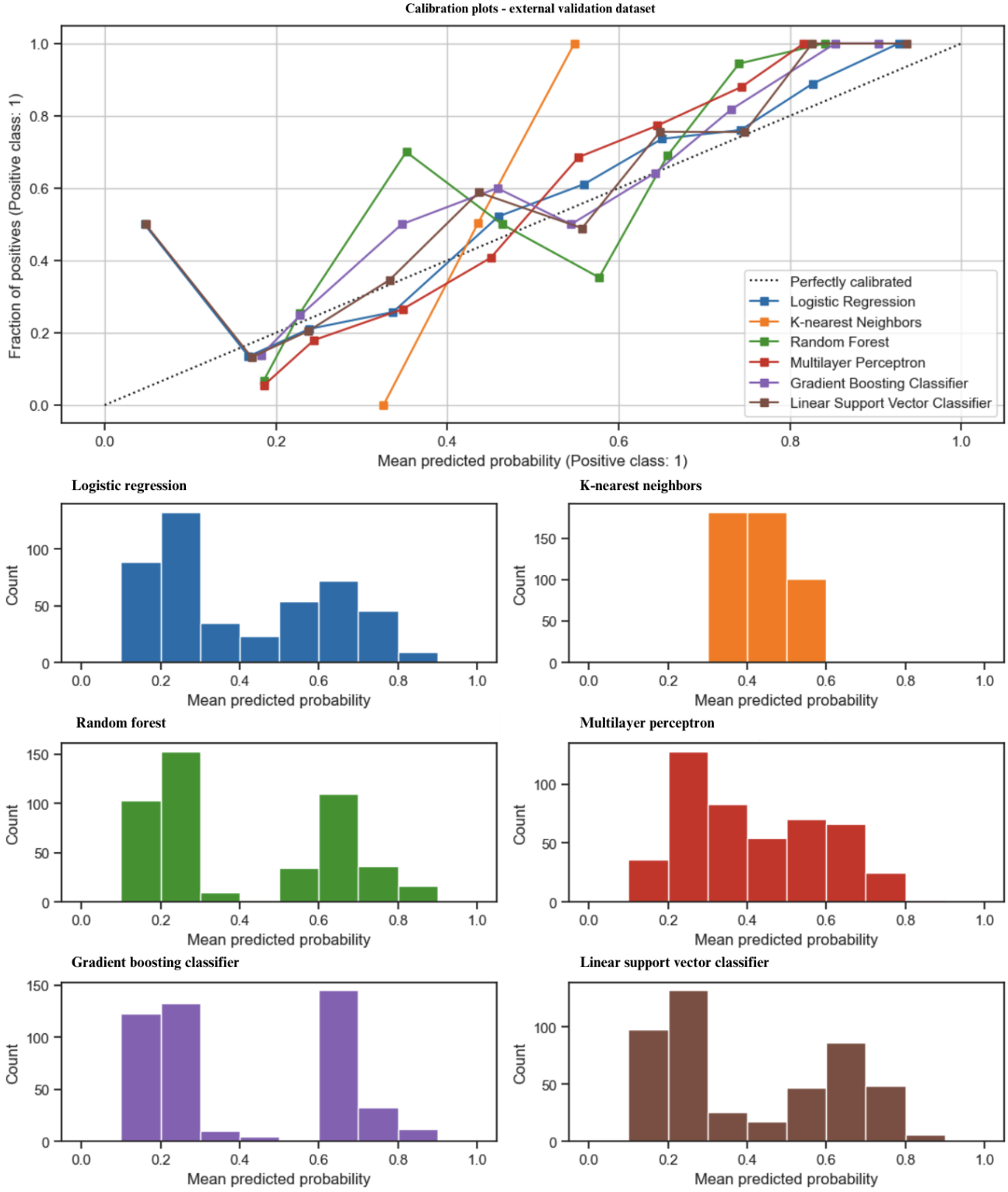
The model calibration was visualized with the calibration plot, which compared the expected probability of frailty, and the mean 10-fold cross-validation predicted the probability of each model. In the internal validation dataset (Figure 2), the LR, SVM, and MLP were well-calibrated with the expected probability of the data. The distribution of predicted probabilities of the models was well-balanced between 0.00 and 1.00. The KNN classifier aligned well with the MPP between 0.00 and 0.40 but it overestimated the high predicted probability. In contrast, the GBC underestimated MPP between 0.20 and 0.40 but aligned well with the rest of the expected probability. These results suggest that LR, SVM, and MLP are the most reliable for balanced predictions, while KNN, GBC, and RF require careful consideration depending on the probability range of interest.

Figure 2. Model calibration of the developed models with internal validation.



In the external validation dataset (Figure 3), the LR, SVM, and MLP were almost perfectly calibrated to the expected probability of the data with a slight overestimation in the 0.40 - 0.80 MPP for MLP and an overestimation in the 0.00 - 0.40 MPP for LR and SVM. These 3 models have a similarly balanced distribution of MPP. RF aligned poorly with an overestimation in the low and high MPP, and it underestimated in the range between 0.40 and 0.70. The GBC aligned well with the calibration plot after 0.60 but overestimated before 0.40. The KNN classifier was poorly calibrated to the plot, giving only MPP in the range of 0.30 - 0.60.

Figure 3. Model calibration of the developed models with external validation.



Discussion

Model Performance

The internal validated models all showed remarkable discrimination performance, with the KNN model having the highest overall discrimination performance. However, the KNN model performed poorly in external validation datasets, indicating overfitting and the inability to generalize its results to other populations. In terms of model optimism, the LR, SVM, and MLP models had better discrimination capabilities, despite a slight decline in the models' performance during external validation. The mean AUCs of these models remained in the range of fair to good performance. When considering all aspects of model performance, the LR model was the most preferable because it provided good discrimination (AUC), an optimal range of predicted probabilities, and good calibration in both the internal and external validation datasets. On the other hand, other models had predicted probabilities that were not extreme enough, resulting in poor to fair calibrations, and they lacked

consistency between the internal and external validation datasets. In comparing the performance and validation aspects of previous frailty prediction models, most studies showed better performance than our best model (LR with a mean AUC of 0.81 in the internal validation set) [45]. The ELSA cohort study (2023) [46] reported the best performance from an RF model with an internal AUC of 0.92. Similarly, Bu et al (2023) [47] presented a multivariate LR model for predicting frailty in diabetic patients that had an AUC of 0.88 in the internal validation set. Additionally, a study on predictive modeling for frailty in older people using various machine learning methods found that artificial neural networks and SVM had the best performance in predicting mortality, with accuracies around 0.78 - 0.79 [32]. Although these studies reported high internal performance metrics, the absence of external validation raises concerns about their generalizability, especially for our focus on the community-dwelling older adult population in general. Our LR showed a mean AUC of 0.75 for the external validation dataset, which dropped slightly from the internal validated results, demonstrating a better edge in the community setting for our focused population. Furthermore, the predictors included in the previously reported model played an important role in prediction performance. Particularly, the utilization of handgrip strength in the study conducted by Bu et al [47] and the assessment of balance and chair stand in the ELSA cohort study [46] could potentially explain the high reported performance in both studies. These predictors either formed part of the frailty phenotype or served as surrogates for physical performance. However, incorporating these predictors into the models necessitated a trade-off between their value added to model performance and the difficulties of model utilization due to the time and skill required for the assessments. In our study, we proposed more parsimonious models using simpler predictions, which still achieved satisfactory performance for frailty screening.

Feature Selection and Findings Explanation

This study's findings suggest that machine learning models can be effective in classifying frailty status among community-dwelling older adults in Northern Thailand using simple predictors including, age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion. Age, BMI, waist circumference, and calf circumference are all potent risk factors for frailty in older adults. When the human body deviates from the normal physiologic process of aging, our levels of estrogen and testosterone gradually decline. These hormones play a vital role in maintaining muscle and bone mass, enhancing strength, and promoting optimal nervous system function [48,49]. As a result, the aging process can accelerate the decline of muscle, bone, and the nervous system, transforming an individual from fit to frail. Additionally, the female population is at higher risk of frailty because the normal bone turnover cycle is disrupted by estrogen deficiency during menopause, increasing bone resorption over deposition, and resulting in net bone loss in women [50]. Household living arrangement was also added to model features as we found that there were studies that showed an association of social adversity and support with frailty status [51,52]. Lastly, we selected level of exhaustion as a predictor in our models, as it had the highest crude aOR among all other features, and we found that it is highly feasible and time-efficient to acquire this data in real clinical settings, compared to other Fried's phenotypes, which involve multiple anthropometric and physical performance tests. These predictors also showed a significant contribution to the frailty prediction in the previous studies [46,47]. Most of our features used in the model are easily collectible, which makes it highly feasible and time-efficient to acquire all data in real clinical settings without depending on a high-level professional, making our model highly applicable for early frailty screening.

Limitations

The most important limitation of this study's models is the generalizability of the models. Our internal and external validation datasets were collected from a community-dwelling population in Lampang and Chiang Mai, respectively, which we confidently believe means that our models have high generalizability to the general population in Thailand. From our perspective, we assume that the models could also be implemented in other Asian countries, as studies showed similar frailty prevalence and population characteristics such as anthropometric measurements, age, household living arrangement, and underlying diseases of hypertension and diabetes [53-58].

However, we encourage conducting validation studies for other regions and populations before clinical application as frailty risk factors do vary across countries as well as in regions with different socioeconomic and health care contexts [59-61]. Another limitation of our study is spectrum bias, as our models were only able to distinguish between frail and robust older adults, despite Fried's criteria having 3 stages of frailty: robust, prefrail, and frail [62]. Nonetheless, we do not believe this bias will significantly affect our study's primary objective, which is the early detection of frailty in older adults to effectively prescribe nutrition and exercise interventions.

Implications

Identifying frailty early helps patient gain access to interventions like nutritional support and exercise programs faster, improving outcomes for older adults by preventing frailty progression, reducing falls, hospitalizations, and mortality, and enhancing quality of life [63,64]. We see 2 possibilities for the practical application of our developed and validated models. One option is to develop a web application that serves as a frailty screening tool that could be self-assessed by individuals or be used in outpatient clinic settings to screen patients. This would help health care providers efficiently identify patients with frailty who require closer monitoring or interventions. We have implemented the validated models to run on our web application, which can be accessed at the link in the Data Availability section.

Another option is to incorporate our machine learning models into electronic medical record or health surveillance systems. The machine learning models could be integrated with the electronic medical record system to provide automated frailty probability scores for each patient. This would enable health

care professionals to identify patients who require closer monitoring or interventions and could help optimize treatment plans for the community-dwelling population at risk of frailty. Furthermore, this could help future researchers retrieve and analyze frailty data from the hospital easier, leading to a better understanding of the factors that contribute to frailty and the development of more effective interventions, which also promotes more efficient use of resources within the health care system.

Conclusion

Machine learning models were fairly good at classifying frailty status among Thai community-dwelling older adults using age, gender, household living arrangement, underlying diseases (hypertension and dyslipidemia), BMI, waist and calf circumference, and level of exhaustion as predictors. The LR and RF models demonstrated the best discrimination performance and model calibration in both the internal and external validation datasets.

There are 2 potential practical applications for utilizing the study findings. These include creating a web application for self-screening or individual screening and incorporating machine learning models into electronic medical record or surveillance systems to provide automated frailty probability scores for individual patients. We advocate for further research on model external validation and temporal recalibration to enhance the model's practicality and applicability to the specific context in which it is used.

Acknowledgments

This work was supported by Chiang Mai University.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The validated model's prediction and application interface can be accessed through the web application [65].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials
[DOCX File, 15 KB - aging_v8i1e62942_app1.docx]

References

1. World population ageing, 2019: highlights. United Nations Digital Library. URL: http://digitallibrary.un.org/record/3846855 [accessed 2025-03-23]
2. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2022-08-02]
3. Population ageing in Thailand. ERIA. URL: http://www.eria.org/publications/population-ageing-in-thailand/ [accessed 2022-08-02]
4. Tantirat P, Suphanchaimat R, Rattanathumsakul T, Noree T. Projection of the number of elderly in different health states in Thailand in the next ten years, 2020-2030. Int J Environ Res Public Health 2020 Nov 23;17(22):8703. [doi: 10.3390/ijerph17228703] [Medline: 33238588]
5. O'Caoimh R, Sezgin D, O'Donovan MR, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. Age Ageing 2021 Jan 8;50(1):96-104. [doi: 10.1093/ageing/afaa219] [Medline: 33068107]
6. Siriwardhana DD, Hardoon S, Rait G, Weerasinghe MC, Walters KR. Prevalence of frailty and prefrailty among community-dwelling older adults in low-income and middle-income countries: a systematic review and meta-analysis. BMJ Open 2018 Mar 1;8(3):e018195. [doi: 10.1136/bmjopen-2017-018195] [Medline: 29496895]
7. Nguyen AT, Nguyen LH, Nguyen TX, et al. Frailty prevalence and association with health-related quality of life impairment among rural community-dwelling older adults in Vietnam. Int J Environ Res Public Health 2019 Oct 12;16(20):3869. [doi: 10.3390/ijerph16203869] [Medline: 31614836]
8. Siriwardhana DD, Weerasinghe MC, Rait G, Falcaro M, Scholes S, Walters KR. Prevalence of frailty in rural community-dwelling older adults in Kegalle district of Sri Lanka: a population-based cross-sectional study. BMJ Open 2019 Jan 25;9(1):e026314. [doi: 10.1136/bmjopen-2018-026314] [Medline: 30782757]
9. Wu C, Smit E, Xue QL, Odden MC. Prevalence and correlates of frailty among community-dwelling Chinese older adults: the China Health and Retirement Longitudinal Study. J Gerontol A Biol Sci Med Sci 2017 Dec 12;73(1):102-108. [doi: 10.1093/gerona/glx098] [Medline: 28525586]

10. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
11. Vaingankar JA, Chong SA, Abdin E, et al. Prevalence of frailty and its association with sociodemographic and clinical characteristics, and resource utilization in a population of Singaporean older adults. Geriatr Gerontol Int 2017 Oct;17(10):1444-1454. [doi: 10.1111/ggi.12891] [Medline: 27576598]
12. Semmarath W, Seesen M, Yodkeeree S, et al. The association between frailty indicators and blood-based biomarkers in early-old community dwellers of Thailand. Int J Environ Res Public Health 2019 Sep 17;16(18):3457. [doi: 10.3390/ijerph16183457] [Medline: 31533354]
13. Srinonprasert V, Chalermsri C, Aekplakorn W. Frailty index to predict all-cause mortality in Thai community-dwelling older population: a result from a National Health Examination Survey cohort. Arch Gerontol Geriatr 2018;77:124-128. [doi: 10.1016/j.archger.2018.05.002] [Medline: 29751290]
14. Xue QL. The frailty syndrome: definition and natural history. Clin Geriatr Med 2011 Feb;27(1):1-15. [doi: 10.1016/j.cger.2010.08.009] [Medline: 21093718]
15. Dent E, Morley JE, Cruz-Jentoft AJ, et al. Physical frailty: ICFSR International Clinical Practice Guidelines for Identification and Management. J Nutr Health Aging 2019;23(9):771-787. [doi: 10.1007/s12603-019-1273-z] [Medline: 31641726]
16. Setiati S, Laksmi PW, Aryana I, et al. Frailty state among Indonesian elderly: prevalence, associated factors, and frailty state transition. BMC Geriatr 2019 Jul 3;19(1):182. [doi: 10.1186/s12877-019-1198-8] [Medline: 31269921]
17. Robertson DA, Savva GM, Kenny RA. Frailty and cognitive impairment--a review of the evidence and causal mechanisms. Ageing Res Rev 2013 Sep;12(4):840-851. [doi: 10.1016/j.arr.2013.06.004] [Medline: 23831959]
18. End of life care in frailty: falls. British Geriatrics Society. URL: https://www.bgs.org.uk/resources/end-of-life-care-in-frailty-falls [accessed 2022-08-02]
19. Cheng MH, Chang SF. Frailty as a risk factor for falls among community dwelling people: evidence from a meta-analysis. J Nurs Scholarsh 2017 Sep;49(5):529-536. [doi: 10.1111/jnu.12322] [Medline: 28755453]
20. Davis-Ajami ML, Chang PS, Wu J. Hospital readmission and mortality associations to frailty in hospitalized patients with coronary heart disease. Aging and Health Research 2021 Dec;1(4):100042. [doi: 10.1016/j.ahr.2021.100042]
21. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. Br J Gen Pract 2019 Jan;69(678):e61-e69. [doi: 10.3399/bjgp18X700241] [Medline: 30510094]
22. Bray NW, Smart RR, Jakobi JM, Jones GR. Exercise prescription to reverse frailty. Appl Physiol Nutr Metab 2016 Oct;41(10):1112-1116. [doi: 10.1139/apnm-2016-0226] [Medline: 27649859]
23. Marcucci M, Damanti S, Germini F, et al. Interventions to prevent, delay or reverse frailty in older people: a journey towards clinical guidelines. BMC Med 2019 Oct 29;17(1):193. [doi: 10.1186/s12916-019-1434-2] [Medline: 31660959]
24. Turner G, Clegg A, British Geriatrics Society, Age UK, Royal College of General Practioners. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age Ageing 2014 Nov;43(6):744-747. [doi: 10.1093/ageing/afu138] [Medline: 25336440]
25. Bergman H, Ferrucci L, Guralnik J, et al. Frailty: an emerging research and clinical paradigm--issues and controversies. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):731-737. [doi: 10.1093/gerona/62.7.731] [Medline: 17634320]
26. Savel TG, Foldy S, Centers for Disease Control and Prevention. The role of public health informatics in enhancing public health surveillance. MMWR Suppl 2012 Jul 27;61(3):20-24. [Medline: 22832993]
27. Gamache R, Kharrazi H, Weiner JP. Public and population health informatics: the bridging of big data to benefit communities. Yearb Med Inform 2018 Aug;27(1):199-206. [doi: 10.1055/s-0038-1667081] [Medline: 30157524]
28. Yang H, Chang J, He W, Wee CF, Yit JST, Feng M. Frailty modeling using machine learning methodologies: a systematic review with discussions on outstanding questions. IEEE J Biomed Health Inform 2024;29(1):631-642. [doi: 10.1109/JBHI.2024.3430226]
29. Bian C, Ye B, Chu CH, McGilton KS, Mihailidis A. Technology for home-based frailty assessment and prediction: a systematic review. Gerontechnology 2020 Nov 2;19(3):1-13. [doi: 10.4017/gt.2020.19.003.06]
30. Mohanty SD, Lekan D, McCoy TP, Jenkins M, Manda P. Machine learning for predicting readmission risk among the frail: explainable AI for healthcare. Patterns (N Y) 2022 Jan 14;3(1):100395. [doi: 10.1016/j.patter.2021.100395] [Medline: 35079714]
31. Aponte-Hao S, Wong ST, Thandi M, et al. Machine learning for identification of frailty in Canadian primary care practices. Int J Popul Data Sci 2021;6(1):1650. [doi: 10.23889/ijpds.v6i1.1650] [Medline: 34541337]
32. Tarekegn A, Ricceri F, Costa G, Ferracin E, Giacobini M. Predictive modeling for frailty conditions in elderly people: machine learning approaches. JMIR Med Inform 2020 Jun 4;8(6):e16678. [doi: 10.2196/16678] [Medline: 32442149]
33. Huang L, Chen H, Liang Z. Enhancing the convenience of frailty index assessment for elderly Chinese people with machine learning methods. Sci Rep 2024 Oct 5;14(1):23227. [doi: 10.1038/s41598-024-74194-x] [Medline: 39369089]
34. Zhang W, Wang J, Xie F, et al. Development and validation of machine learning models to predict frailty risk for elderly. J Adv Nurs 2024 Dec;80(12):5064-5075. [doi: 10.1111/jan.16192] [Medline: 38605460]
35. Seesen M, Sirikul W, Ruangsuriya J, Griffiths J, Siviroj P. Cognitive frailty in Thai community-dwelling elderly: prevalence and its association with malnutrition. Nutrients 2021 Nov 25;13(12):4239. [doi: 10.3390/nu13124239] [Medline: 34959791]

36. Thinuan P, Siviroj P, Lerttrakarnnon P, Lorga T. Prevalence and potential predictors of frailty among community-dwelling older persons in Northern Thailand: a cross-sectional study. Int J Environ Res Public Health 2020 Jun 8;17(11):4077. [doi: 10.3390/ijerph17114077] [Medline: 32521642]
37. Haque MN. Active ageing level of older persons: regional comparison in Thailand. J Aging Res 2016;2016:9093018. [doi: 10.1155/2016/9093018] [Medline: 27375903]
38. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
39. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
40. Riley RD, Snell KIE, Archer L, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ 2024 Jan 22;384:e074821. [doi: 10.1136/bmj-2023-074821] [Medline: 38253388]
41. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. J Mach Learn Res 2017;18(17):1-5.
42. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. jair 2002 Jun 1;16:321-357. [doi: 10.1613/jair.953]
43. Chittrakul J, Siviroj P, Sungkarat S, Sapbamrer R. Physical frailty and fall risk in community-dwelling older adults: a cross-sectional study. J Aging Res 2020;2020:3964973. [doi: 10.1155/2020/3964973] [Medline: 32714612]
44. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
45. Velazquez-Diaz D, Arco JE, Ortiz A, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: scoping review. J Med Internet Res 2023 Oct 20;25(1):e47346. [doi: 10.2196/47346] [Medline: 37862082]
46. Leme DDC, de Oliveira C. Machine learning models to predict future frailty in community-dwelling middle-aged and older adults: the ELSA cohort study. J Gerontol A Biol Sci Med Sci 2023 Oct 28;78(11):2176-2184. [doi: 10.1093/gerona/glad127] [Medline: 37209408]
47. Bu F, Deng XH, Zhan NN, et al. Development and validation of a risk prediction model for frailty in patients with diabetes. BMC Geriatr 2023 Mar 27;23(1):172. [doi: 10.1186/s12877-023-03823-3] [Medline: 36973658]
48. Dobbs MB, Buckwalter J, Saltzman C. Osteoporosis: the increasing role of the orthopaedist. Iowa Orthop J 1999;19:43-52. [Medline: 10847516]
49. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jul 1;48(4):601-601. [doi: 10.1093/ageing/afz046] [Medline: 31081853]
50. Truumees E. Chapter 22 - Osteoporosis. In: Errico TJ, Lonner BS, Moulton AW, editors. Surg Manag Spinal Deform: W.B. Saunders; 2009:351-372. [doi: 10.1016/B978-141603372-1.50025-1]
51. Gomes CDS, Guerra RO, Wu YY, et al. Social and economic predictors of worse frailty status occurrence across selected countries in North and South America and Europe. Innov Aging 2018 Sep;2(3):igy037. [doi: 10.1093/geroni/igy037] [Medline: 30569024]
52. Lee Y, Chon D, Kim J, Ki S, Yun J. The predictive value of social frailty on adverse outcomes in older adults living in the community. J Am Med Dir Assoc 2020 Oct;21(10):1464-1469. [doi: 10.1016/j.jamda.2020.03.010] [Medline: 32362535]
53. Yu R, Wu WC, Leung J, Hu SC, Woo J. Frailty and its contributory factors in older adults: a comparison of two Asian regions (Hong Kong and Taiwan). Int J Environ Res Public Health 2017 Sep 21;14(10):1096. [doi: 10.3390/ijerph14101096] [Medline: 28934150]
54. Liu ZY, Wei YZ, Wei LQ, et al. Frailty transitions and types of death in Chinese older adults: a population-based cohort study. Clin Interv Aging 2018;13:947-956. [doi: 10.2147/CIA.S157089] [Medline: 29805253]
55. To TL, Doan TN, Ho WC, Liao WC. Prevalence of frailty among community-dwelling older adults in Asian countries: a systematic review and meta-analysis. Healthcare (Basel) 2022 May 12;10(5):895. [doi: 10.3390/healthcare10050895] [Medline: 35628034]
56. Huynh TQH, Pham TLA, Vo VT, Than HNT, Nguyen TV. Frailty and associated factors among the elderly in Vietnam: a cross-sectional study. Geriatrics (Basel) 2022 Aug 20;7(4):85. [doi: 10.3390/geriatrics7040085] [Medline: 36005261]
57. Norazman CW, Adznam SN, Jamaluddin R. Physical frailty among urban-living community-dwelling older adults in Malaysia. Int J Environ Res Public Health 2020 Sep 9;17(18):6549. [doi: 10.3390/ijerph17186549] [Medline: 32916809]
58. Murayama H, Kobayashi E, Okamoto S, et al. National prevalence of frailty in the older Japanese population: findings from a nationally representative survey. Arch Gerontol Geriatr 2020;91:104220. [doi: 10.1016/j.archger.2020.104220] [Medline: 32810734]
59. Morkphrom E, Srinonprasert V, Sura-amonrattana U, et al. Severity of frailty using modified Thai frailty index, social factors, and prediction of mortality among community-dwelling older adults in a middle-income country. Front Med (Lausanne) 2022;9:1060990. [doi: 10.3389/fmed.2022.1060990] [Medline: 36569139]
60. Qin Y, Hao X, Lv M, Zhao X, Wu S, Li K. A global perspective on risk factors for frailty in community-dwelling older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr 2023 Feb;105:104844. [doi: 10.1016/j.archger.2022.104844] [Medline: 36335672]

61. Zeng XZ, Meng LB, Li YY, et al. Prevalence and factors associated with frailty and pre-frailty in the older adults in China: a national cross-sectional study. Front Public Health 2023;11:1110648. [doi: 10.3389/fpubh.2023.1110648] [Medline: 37554734]
62. Willis BH. Spectrum bias--why clinicians need to be cautious when applying diagnostic test studies. Fam Pract 2008 Oct;25(5):390-396. [doi: 10.1093/fampra/cmn051] [Medline: 18765409]
63. Rockwood K, Song X, Mitnitski A. Changes in relative fitness and frailty across the adult lifespan: evidence from the Canadian National Population Health Survey. CMAJ 2011 May 17;183(8):E487-E494. [doi: 10.1503/cmaj.101271] [Medline: 21540166]
64. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
65. Frailty Classificiation Using Machine Learning Model. URL: https://frailty-commed-cmu.streamlit.app [accessed 2025-03-23]

Abbreviations

aOR: adjusted odds ratio
AUC: area under the receiver operating characteristic curve
CC: calf circumference
GBC: gradient boosting classifier
KNN: k-nearest neighbors
LR: logistic regression
MLP: multilayer perceptron artificial neural network
MPP: mean predicted probability
NPV: negative predictive value
PPV: positive predictive value
RF: random forest
SMOTE: synthetic minority oversampling technique
SVM: linear support vector machine classifier
TRIPOD+AI: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence

Edited by J Wang; submitted 05.06.24; peer-reviewed by Y Huang, Y Huusi; revised version received 26.11.24; accepted 28.02.25; published 02.04.25.

Please cite as:
Isaradech N, Sirikul W, Buawangpong N, Siviroj P, Kitro A
Machine Learning Models for Frailty Classification of Older Adults in Northern Thailand: Model Development and Validation Study
JMIR Aging 2025;8:e62942
URL: https://aging.jmir.org/2025/1/e62942
doi:10.2196/62942

Identification of Target Body Composition Parameters by Dual-Energy X-Ray Absorptiometry, Bioelectrical Impedance, and Ultrasonography to Detect Older Adults With Frailty and Prefrailty Status Using a Mobile App in Primary Care Services: Descriptive Cross-Sectional Study

Beatriz Ortiz-Navarro[1*], MD; José Losa-Reyna[2,3*], PhD; Veronica Mihaiescu-Ion[4,5*], PT, MSc; Jerónimo Garcia-Romero[6*], MD, Prof Dr; Margarita Carrillo de Albornoz-Gil[6*], MD, PhD; Alejandro Galán-Mercant[4,5*], PhD, PT, Prof Dr

[1]Public Andalusian Health System, Málaga, Spain

[2]CIBER of Frailty and Healthy Aging, Madrid, Spain

[3]Sports Science Research Centre, King Juan Carlos University, Madrid, Spain

[4]Institute of Biomedicine of Cádiz, Cádiz, Spain

[5]MOVE-IT Research Group, Department of Nursing and Physiotherapy, Faculty of Health Sciences, University of Cádiz, Cádiz, Spain

[6]Área de Educación Física y Deportiva, Facultad de Medicina, Universidad de Málaga, Boulevard Louis Pasteur, Málaga, Spain

[*]all authors contributed equally

Corresponding Author:
Margarita Carrillo de Albornoz-Gil, MD, PhD
Área de Educación Física y Deportiva, Facultad de Medicina, Universidad de Málaga, Boulevard Louis Pasteur, Málaga, Spain

Abstract

Background: Frailty syndrome in older adults represents a significant public health concern, characterized by a reduction in physiological reserves and an increased susceptibility to stressors. This can result in adverse health outcomes, including falls, hospitalization, disability, and mortality. The early identification and management of frailty are essential for improving quality of life and reducing health care costs. Conventional assessment techniques, including dual-energy X-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and muscle ultrasound (US), are efficacious but frequently constrained in primary care settings by financial and accessibility limitations.

Objective: The aim of this study is to analyze the differences in anthropometric characteristics, physical function, nutritional status, cognitive status, and body composition among older adults identified as frail, prefrail, or robust in primary care services using the PowerFrail mobile app. Furthermore, the study assesses the predictive capacity of body composition variables (whole-body phase angle [WBPhA] via BIA, US-measured rectus femoris muscle thickness, and DXA-derived lean mass) in identifying frailty and evaluates their feasibility for implementation in primary care.

Methods: A descriptive cross-sectional study was conducted with 94 older adult participants aged between 70 and 80 years, recruited through the Andalusian Health Service in Spain. Frailty status was classified using the PowerFrail App, which integrates muscle power assessment and provides personalized physical activity recommendations. Body composition was measured using WBPhA (BIA), muscle US, and DXA. Statistical analyses included 1-way ANOVA for group comparisons, logistic regression to investigate associations, and receiver operating characteristic curve analysis to evaluate the predictive accuracy of the body composition measures.

Results: Participants were categorized into frail (n=28), prefrail (n=33), and robust (n=33) groups. All body composition measures exhibited high specificity in detecting frailty, with varying sensitivity. Unadjusted US showed the highest specificity but low sensitivity (10.7%). WBPhA and right leg lean mass (LeanM RL) demonstrated significant predictive capabilities, especially when adjusted for age and sex, with area under the curve values ranging from 0.678 to 0.762. The adjusted LeanM RL model showed a good balance between sensitivity (35.7%) and specificity (93.9%; *P*=.045), indicating its potential as a reliable frailty predictor. These findings are consistent with previous research emphasizing the importance of muscle mass and cellular health in frailty assessment.

Conclusions: Body composition variables, particularly WBPhA, LeanM RL, and US, are effective predictors of frailty in older adults. The PowerFrail mobile app, combined with advanced body composition analysis, offers a practical and noninvasive method

for early frailty detection in primary care settings. Integrating such technological tools can enhance the early identification and management of frailty, thereby improving health outcomes in the aging population.

(JMIR Aging 2025;8:e67982) doi:10.2196/67982

KEYWORDS

frailty syndrome; older adults; body composition; bioelectrical impedance analysis; muscle ultrasound, dual-energy X-Ray absorptiometry; mobile health apps; primary care; PowerFrail App

Introduction

Frailty syndrome in older adults is a multifaceted clinical condition characterized by decreased physiological reserves and increased vulnerability to stressors, which elevates the risk of adverse health outcomes such as falls, hospitalization, disability, and mortality [1,2]. As the global population ages, the prevalence of frailty is projected to rise, posing significant challenges to health care systems worldwide [3]. Early identification and management of frailty are crucial for enhancing quality of life, maintaining functional independence, and reducing health care costs associated with frailty-related complications [4,5].

Primary care facilities are uniquely positioned to play a pivotal role in the early detection and management of frailty due to their accessibility and continuous engagement with the aging population [6]. Implementing effective screening tools within primary care can facilitate timely interventions, thereby mitigating the progression of frailty and its associated adverse outcomes [7]. However, the integration of comprehensive frailty assessments into routine primary care practice remains limited, often due to time constraints, lack of standardized tools, and insufficient training among primary care providers [8].

Body composition variables, particularly those related to muscle mass and tissue quality, are essential in identifying and predicting frailty syndrome [9]. Techniques such as dual-energy X-ray absorptiometry (DXA), bioelectrical impedance analysis (BIA), and muscle ultrasound (US) have been extensively used to assess body composition and muscle status in older adults [10]. The whole-body phase angle (WBPhA), obtained through BIA, serves as an indicator of cellular health and nutritional status, with lower values associated with increased frailty and poorer clinical outcomes [11,12]. Similarly, muscle US offers a noninvasive way to assess muscle thickness and quality, aiding in the detection of sarcopenia and frailty [13]. Although DXA is considered the gold standard for measuring bone and muscle mass, its high cost and limited accessibility in primary care facilities make it necessary to explore alternative assessment methods [14].

In recent years, the advent of mobile health (mHealth) apps has introduced innovative solutions for the assessment and monitoring of complex geriatric syndromes, including frailty. The PowerFrail app represents a pioneering effort in this domain, being the first clinical and scientifically validated app designed to assess muscle power and frailty in older adults in a user-friendly manner [15,16]. This app not only facilitates the screening process but also provides individualized recommendations for improvement and tailored physical activity regimens, thereby supporting personalized intervention strategies. The utilization of validated mobile apps and new trends around artificial intelligence in primary care hold significant promise for enhancing the early detection and management of frailty, offering advantages such as accessibility, ease of use, and the ability to provide real-time feedback and recommendations [17,18].

Moreover, integrating mHealth tools into primary care can bridge gaps in health care delivery by enabling continuous monitoring and follow-up, which are critical for managing chronic conditions and preventing the escalation of frailty [18]. Previous studies have highlighted the effectiveness of mobile apps in improving health outcomes among older adults by facilitating timely interventions and enhancing patient engagement [19]. These findings underscore the potential of mHealth solutions to complement traditional assessment methods, providing a comprehensive approach to frailty management in primary care settings.

From above, the main objective of this study is to analyze the differences in variables related to anthropometric characteristics, physical function, nutritional status, cognitive status, and body composition in phenotypes of frail, prefrail, and robust older adults identified in primary care services. Frailty levels will be classified using the PowerFrail app. Additionally, the second objective is to assess the predictive capacity of body composition variables (US, bioimpedance, and DXA) in identifying older adults with frailty and to evaluate their implementation in primary care services. We used the technological tools WBPhA, US for rectus femoris muscle thickness, and DXA for bone mineral density and lean mass assessment. Our results could demonstrate that these parameters are useful predictors for the identification of frailty, in line with previous findings, thus supporting the potential integration of these tools into primary care practices for the early detection and management of frailty, integrating for the first time mHealth technologies with advanced body composition analysis systems.

Methods

Ethical Considerations

This was a descriptive cross-sectional study that evaluated clinical, physiological, body composition, and psychometric variables in a sample of older adult participants. Recruitment and data collection took place between March and June of 2023. Participants were recruited through advertisements from the public Andalusian Health Service system, Spain. Participants were recruited from primary care centers through referrals by health care professionals. Recruitment was conducted in collaboration with general practitioners, who identified potential

participants meeting the inclusion criteria. The study adhered to the ethical principles of the Declaration of Helsinki for medical research involving human subjects. All participants provided written informed consent. This study was approved by the Costa del Sol Institutional Ethics Committee, under protocol number BON22, on December 23, 2022. All participants provided written informed consent prior to their inclusion in the study. Participants' privacy and confidentiality were ensured throughout the study: all data were anonymized prior to analysis, and identifying information was stored securely and separately from the study data. No identifiable images or information of participants are included in this publication. Participants did not receive compensation for their participation..

Participants

Inclusion criteria for the study were as follows:

- Individuals aged between 70 and 80 years, inclusive.
- Absence or presence of mild cognitive impairment as determined by a Barthel index score [20] greater than 95.
- Sit-to-stand test score [21] between 2.5 and 3.6 for men or between 1.9 and 3 for women.
- Signed informed consent.

Exclusion criteria were the following:

- Individuals younger than 70 years or older than 80 years.
- Individuals with moderate or severe cognitive impairment.
- Sit-to-stand test score [21] less than 2.5 or greater than 3.6 for men or less than 1.9 or greater than 3 for women.
- Individuals living in institutions.
- Individuals with pacemakers or metal prostheses.

Variables and Procedure

The aim of this study was to identify differences in measured parameters to classify individuals into different states of frailty within primary care. To identify different states of frailty, the sit-to-stand test was used, and relative power was calculated using the equations validated by Losa-Reyna et al [16]. Using the cutoff points determined by Losa-Reyna et al [22], participants were classified into 3 groups: frail, prefrail, and robust. This test was performed in the health center's office.

After identifying the study subjects, a detailed evaluation of the remaining variables was carried out in the Functional Testing Laboratory of the Physical Education and Sports Area of the University of Málaga. Participants came to the laboratory fasting (a minimum of 3 hours) and without having done any previous exercise, wearing comfortable and light clothing without metal objects. The tests were carried out between 9:30 AM and 2:00 PM. Once in the laboratory, 2 qualified researchers administered the different tests and assessments; body composition determinations were performed fasting and in the early morning, and the rest of the scheduled evaluations were carried out after breakfast.

Clinical and Demographic Variables

We collected the following clinical and demographic variables:

- Age and sex: Recorded during the initial interview (age in years; sex as male or female).
- Body mass index: Calculated by dividing weight (measured in kilograms using a Digital Scale Extra Large Seca Robusta 813) by height (measured in meters using a stadiometer) squared (m^2).
- Number of medications (drugs): Extracted from the clinical history of the Diraya system (Andalusian Health Service database), with the participants' consent.
- Education level: Recorded during the initial interview.

Body Composition Variables

We collected the following body composition variables:

- Muscle architecture: Muscle thickness and pennation angle of the rectus femoris muscle of the dominant leg were measured using US (Logiq Book XP Ultrasound System and 8L linear transducer). The rectus femoris was examined with the participant in a supine position, with the US operator standing on the ipsilateral side of the participant. The evaluation protocol previously described by Mateos-Angulo et al [23] was used.
- DXA: Lean mass of arms and legs, total lean body mass, skeletal muscle index, and bone mineral density were determined using a Hologic Horizon A DXA scanner (Hologic Inc). Each subject was examined by a certified technician. The distinction between bone and soft tissue, the detection of edges, and regional demarcations were performed using computer algorithms with APEX Corporation Software (version 5.6.0.7). For each scan, patients were asked to remove all materials that could attenuate the X-ray beam, including jewelry. Due to the sensitivity of the soft tissue analysis, the patient should only wear a paper gown for the scan. There should be no pillow on the scan, as the material would affect the soft tissue measurement. The densitometer calibration was checked daily with the standard calibration block supplied by the manufacturer.
- BIA: Total body water, intracellular, extracellular water, and phase angle (WBPhA) variables were determined by multifrequency bioimpedance using the Inbody 770 model. Multifrequency segmental data were obtained that accurately determined total body water, intracellular and extracellular water, impedance (Xc and R), and phase angle (Z) in the 5 body segments (right arm, left arm, trunk, right leg, and left leg).

Cognitive and Nutritional Status

To assess cognitive and nutritional status, the following questionnaires were administered digitally using the Google Forms application.

- Cognitive capacity: Cognitive status was evaluated [24] using the General Practitioner Assessment of Cognition, score 0 to 8, a rapid, reliable, and specific test for the detection of dementia in primary care. This instrument is considered an efficient alternative to others, such as the Mini-Mental State Examination, due to its rapid administration and lack of bias related to gender, education level, or mental health.
- Nutritional status screening: Nutritional status was assessed using the Mini Nutritional Assessment (MNA) [25], score

0 to 14, a widely used scale in the geriatric population. The first part of this test serves as a screening tool to detect the risk of malnutrition, with a cutoff value of 10 points or less. It is a widely used method in older adults and has been validated in different clinical contexts.

Physical Function Evaluation

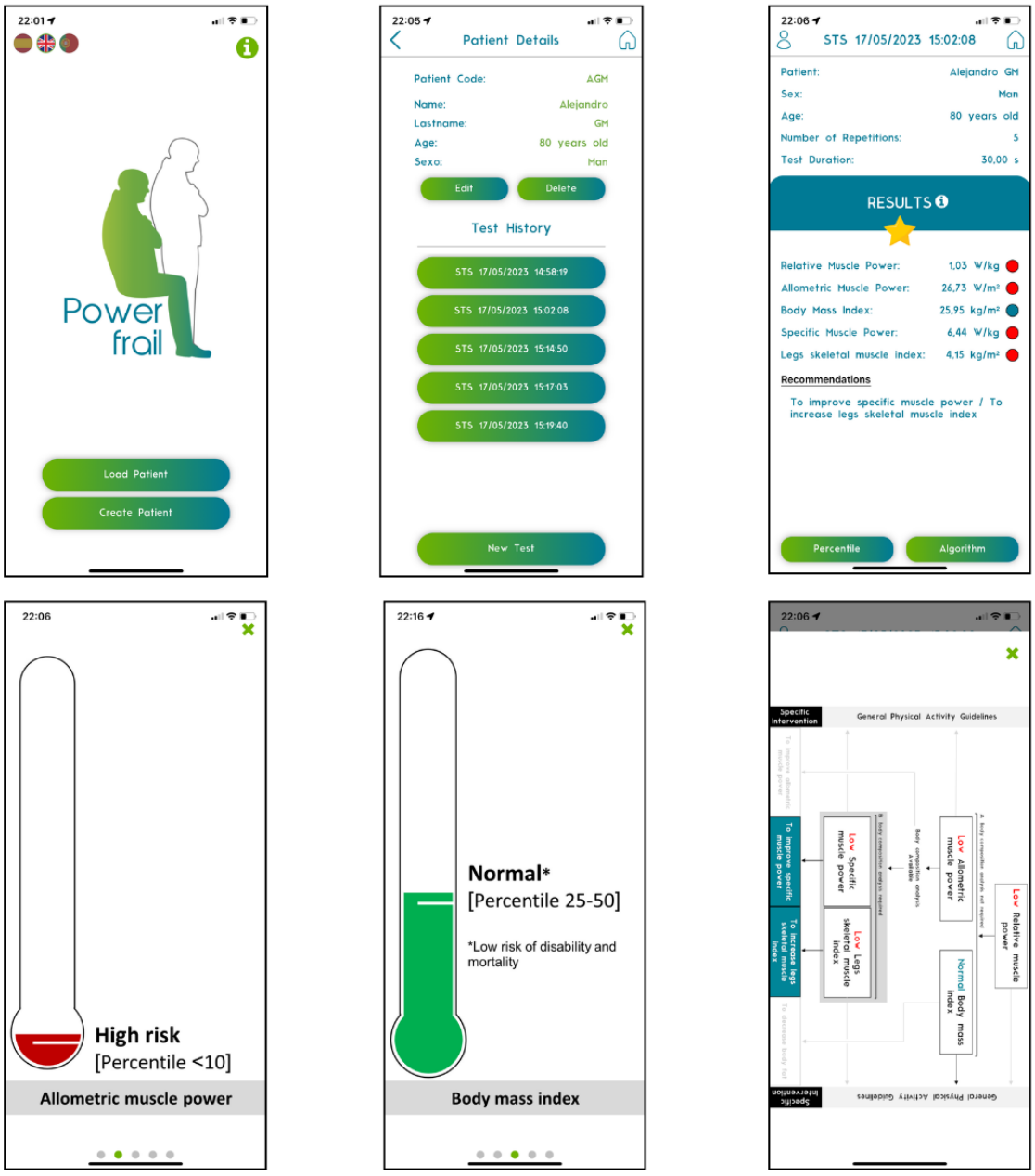
To evaluate the physical function of the participants, the Short Physical Performance Battery (SPPB) was used, score 0 to 12 [26]. This battery includes three tests:

- Walking speed: Measured over a 4-meter distance, expressed in meters per second. Given that gait speed measurement is used as a tool for detecting frailty in older adults from the general population due to its high sensitivity, simplicity, and feasibility, we also used the cutoff points for frailty and prefrailty related to gait speed and sarcopenia in accordance with the consensus document developed by Cruz-Jentoft et al [27]. In this framework, values below 0.6 m/s are identified as frail, between 0.6 and 1 m/s as prefrail, and above 1 m/s as robust [27].
- Static balance: Assessed in 3 different positions.
- Sit-to-stand test: Evaluation of the time it takes for participants to stand up and sit down from a chair 5 times.

In addition to the tests included in the SPPB, two tests were performed:

- Mobile lower limb relative muscle power (RPOW): To measure RPOW, participants performed 5 repetitions of standing up from and sitting down onto a chair with a height of 0.46 m, following the protocol validated by Alcazar et al [15]. The test was performed using the PowerFrail app (Figure 1), developed and validated by Losa-Reyna et al [16], installed on a stable smartphone.
- Handgrip strength: Isometric handgrip strength, expressed in kilograms, was measured using a Takei Physical Fitness Test adjustable dynamometer, following a standardized protocol, as indicated by Roberts et al [28]. Participants performed the test in an erect standing position, with shoulders adducted and arms extended parallel to the body, without touching their torso. Two attempts were made for each extremity, and the maximum value was considered, regardless of hand dominance.

Figure 1. PowerFrail mobile app screenshots.



Sample Size Calculation

The sample size was estimated using an ANOVA model for 3 independent groups (robust, prefrail, and frail). The calculation was performed using the G*Power software. For the sample size calculation, a power of 90% (1–β=.90), a significance level of 5% (α=.05), and an effect size of 0.4 were used. The result was a total sample size of 84 participants. This size was obtained with a noncentrality parameter (λ) of 13.44 and a critical *F* of 3.11, with 2 degrees of freedom in the numerator and 81 in the denominator. To avoid losses in the initial calculation, enough participants were recruited to maintain the robustness of the study. Finally, 94 subjects participated in the study, distributed across 3 groups: robust, prefrail, and frail.

Statistical Analysis

To compare the differences on all variables between robust participants and those with frailty and prefrailty, a 1-way ANOVA was performed on the total sample. Logistic regression analysis was performed to investigate the relationship between body composition and frailty. WBPhA, thickness US, and DXA lean muscle from the lower limbs were entered into the regression model as independent variables, as they were found to have significant differences between the studied groups. Considering that age and sex may influence the relationship between body composition and frailty, these factors were introduced as a confounding variable into an adjusted regression model, treating age as a continuous covariate and sex as a categorical factor. The dependent variable was a binary indicator of frailty, coded as 1 for participants with frailty and 0 for robust participants and those with prefrailty.

To extract values for body composition to identify the presence of prefrailty or frailty, we conducted an analysis using the receiver operating characteristic (ROC) curve. This analysis was applied only to body composition that was significant in the logistic regression analysis. In the ROC analysis, the outcome variable was the presence or absence of frailty. The test variable was the body composition that was significantly associated with frailty. Youden index [29] was calculated with the following formula: Youden index = sensitivity + specificity – 1. The area under the curve (AUC), sensitivity, and specificity were calculated to evaluate the accuracy of the identified predictive models. The AUC could distinguish between nonpredictive (AUC<0.5), less predictive (0.5<AUC<0.7), moderately predictive (0.7<AUC<0.9), and highly predictive (0.9<AUC<1) values, as well as perfect prediction (AUC=1) [30].

Results

The background information of the participants is shown in Table 1. Of the 94 participants in this study, 28 were frail, 33 were prefrail, and 33 had a robust profile. The average age for each frailty category was 76.5, 75.3, and 74.0 years, respectively. The percentages of men and women with frailty and prefrailty were 32% and 52% for men and 68% and 48% for women, respectively. All investigated variables were significantly different among the frailty categories, except BMI, MNA score, skeletal mass index, left arm lean mass, right arm lean mass, and total body lean mass.

Table . Descriptive characteristics of the sample (n=94).

	Frailty[a]	Prefrailty	Robust	*P* value[b]	Post hoc[a]
Particpants, n (%)	28 (30)	33 (35)	33 (35)		
Sex (female; male), n (%)	19 (68); 9 (32)	16 (48); 17 (52)	17 (52); 16 (48)		
Age (years), mean (SD)	76.46 (2.92)	75.34 (2.86)	73.99 (2.59)	.004	*F*>R
BMI (kg/m^2), mean (SD)	28.35 (5.34)	27.87 (4.73)	28.65 (5.27)	.82	
RPOW[c] (W/kg), mean (SD)	1.69 (0.54)	2.60 (0.32)	3.32 (0.46)	<.001	*F*<P<R
MNA[d] (score, 0 - 14), mean (SD)	13.00 (1.28)	12.94 (1.71)	13.30 (0.98)	.52	
GPCOG[e] (score, 0 - 8), mean (SD)	5.46 (1.73)	6.60 (1.32)	7.15 (1.00)	<.001	*F*<P,R
SPPB[f] (score, 0 - 12), mean (SD)	8.07 (2.26)	10.79 (1.24)	11.75 (0.43)	<.001	*F*<P<R
5STS[g] (s), mean (SD)	18.47 (6.44)	12.01 (1.40)	9.27 (1.31)	<.001	*F*<P<R
WALK4m[h] (s), mean (SD)	5.92 (2.25)	4.07 (0.75)	3.92 (0.61)	<.001	*F*<P<R
GS4m[i] (m/s), mean (SD)	0.74 (0.21)	1.01 (0.16)	1.04 (0.15)	<.001	*F*<P,R
Handgrip max (kg), mean (SD)	19.24 (5.39)	25.48 (6.57)	26.25 (7.50)	<.001	*F*<P,R
Number of drugs[j], mean (SD)	6.89 (2.81)	4.73 (3.13)	3.09 (2.20)	<.001	*F*>P>R
US[k] (cm), mean (SD)	0.92 (0.32)	1.02 (0.25)	1.16 (0.26)	.004	*F*<R
WBPhA[l] (°), mean (SD)	4.36 (0.55)	4.63 (0.61)	4.95 (0.43)	<.001	*F*<R
SMI[m], mean (SD)	6.51 (1.04)	7.09 (1.08)	7.20 (1.33)	.07	
Left arm lean mass (g), mean (SD)	2088.24 (624.67)	2485.68 (720.65)	2408.92 (772.28)	.08	
Right arm lean mass (g), mean (SD)	2274.00 (597.82)	2624.04 (787.75)	2661.02 (915.64)	.12	
Left leg lean mass (g), mean (SD)	5975.11 (1494.27)	7147.72 (1794.43)	7158.07 (2192.25)	.02	*F*<P,R
Right leg lean mass (g), mean (SD)	5982.38 (1573.34)	7368.42 (1823.40)	7314.78 (2040.34)	.006	*F*<P,R
Total body lean mass (g), mean (SD)	42,130.41 (8174.96)	48,123.28 (11,063.12)	47,858.20 (12,130.27)	.06	

[a]F: frailty; P: prefrailty; R: robust.

[b]*P* value adjusted for comparing a family of 3.

[c]RPOW: relative muscle power.

[d]MNA: Mini Nutritional Assessment.

[e]GPCOG: General Practitioner Assessment of Cognition.

[f]SPPB: Short Physical Performance Battery.

[g]5STS: 5 sit-to-stand time.

[h]WALK4m: time to walk 4 meters.

[i]GS4m: gait speed over 4 meters at a normal pace.

[j]Drugs: number of daily medications.

[k]US: ultrasound on rectus femoris.

[l]WBPhA: whole-body phase angle.

[m]SMI: skeletal muscle index.

The correlation analysis is presented through a heat map in Figures 2-4. Starting with the primary variable identified as a frailty criterion (RPOW), we initially analyzed the behavior of this variable in relation to the other studied variables. Additional correlation analyses were also explored, such as examining the behavior and relationship of handgrip strength. In this regard, we conducted a first analysis on the entire sample (Figure 2), then one each for women (Figure 3) and men (Figure 4). In the entire sample (Figure 1), we found strong positive and negative correlations between the RPOW and all the variables, except MNA score, however, these relationships changed for the women-only and men-only samples. In the complete sample (Figure 2), the results identified stronger correlations between SPPB and RPOW, 5 sit-to-stand time and RPOW, drugs and RPOW, GS4m and RPOW, and handgrip and RPOW. Regarding the analysis of the correlations separately for women and men (Figures 3 and 4), the results were as expected. Specifically, as observed in the cited figures, there were relationships involving body composition variables (BIA, DXA, and US) and expressions of strength, such as lower limb power and, more markedly, handgrip strength. The observed differences in correlations between sexes are consistent with expectations, given the complex biological, hormonal, and social interactions that affect males and females differently during the aging process. It is important to consider these factors when interpreting the results and when designing interventions that address frailty and functional decline in older adults effectively and equitably.

Figure 2. Simple correlation analysis results for the whole sample (n=94). Pearson correlations of all studied variable were carried out on the total sample. 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.

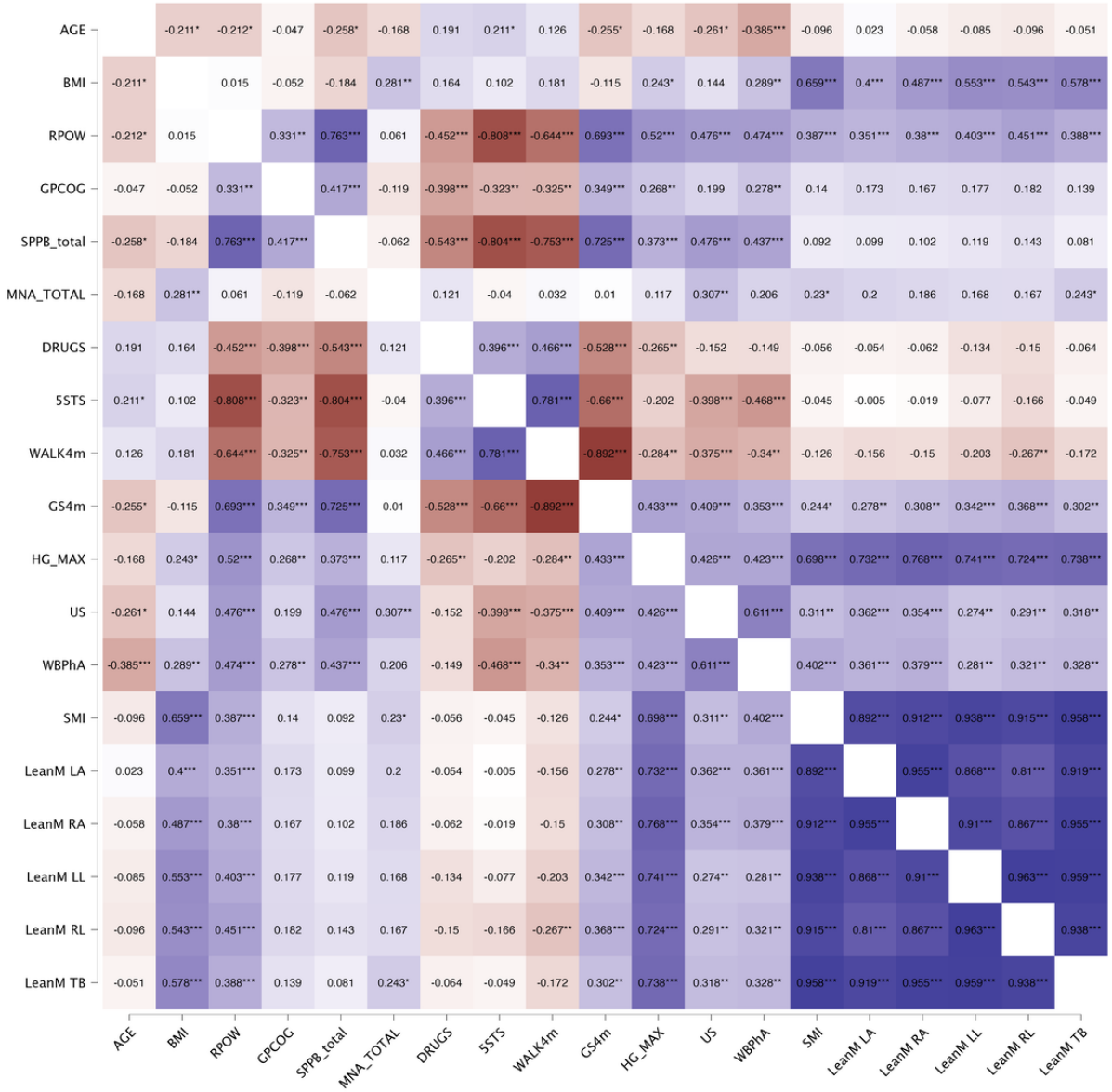


Figure 3. Simple correlation analysis results for women (n=52). Pearson correlations were conducted for all studied variables on the sample of women. 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.

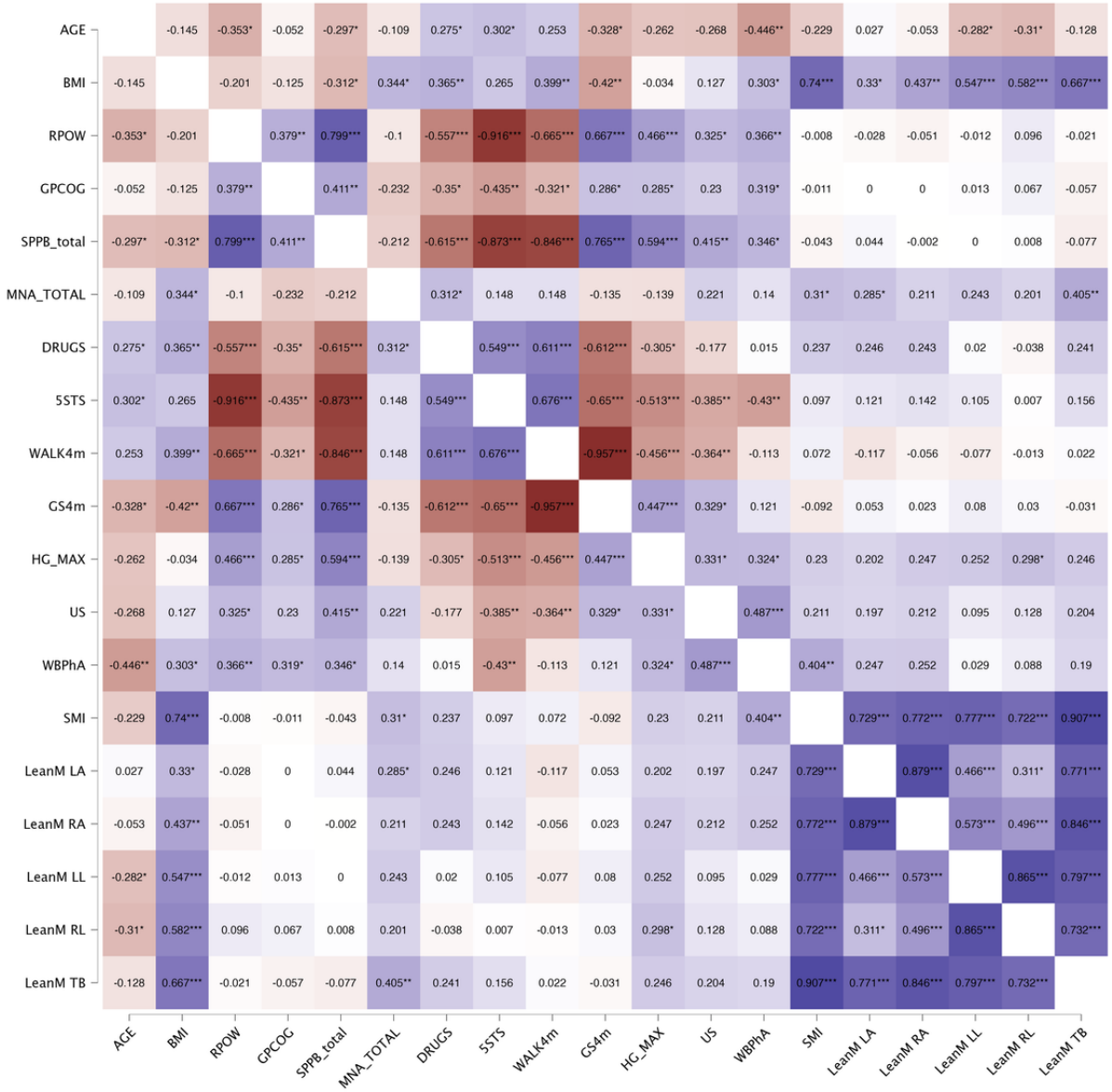
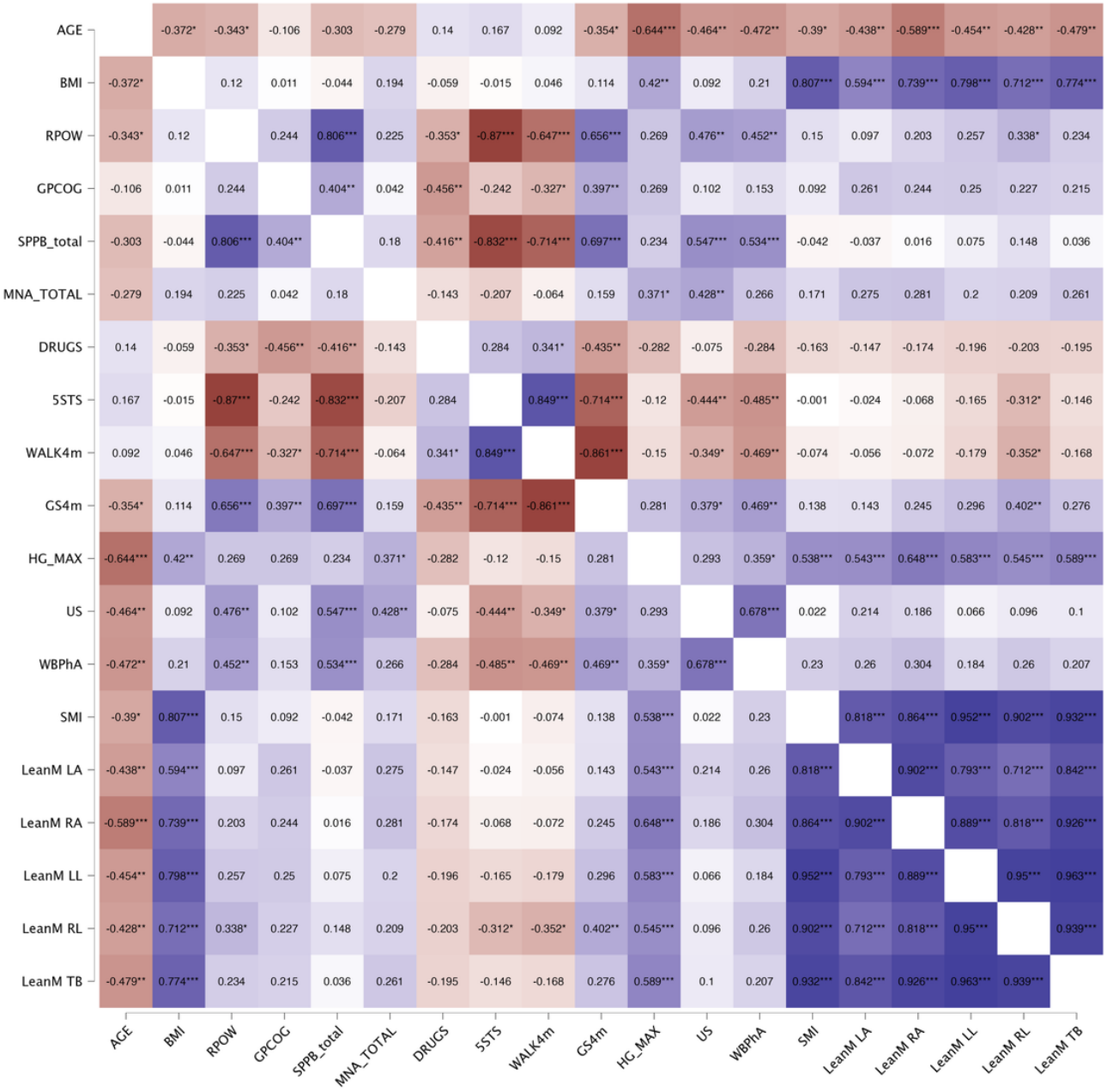


Figure 4. Simple correlation analysis results for men (n=42). Pearson correlations were conducted for all studied variables on the sample of men (n=42). 5STS: 5 sit-to-stand time; DRUGS: number of daily medications; GPCOG: General Practitioner Assessment of Cognition score; GS4m: gait speed over 4 meters at a normal pace; HG: handgrip; LeanM LA: left arm lean mass; LeanM LL: left leg lean mass; LeanM RA: right arm lean mass; LeanM RL: right leg lean mass; LeanM TB: total body lean mass; MNA: Mini Nutritional Assessment score; RPOW: relative muscle power; SMI: skeletal muscle index; SPPB: Short Physical Performance Battery score; US: ultrasound on rectus femoris; WALK4m: time to walk 4 meters; WBPhA: whole-body phase angle. *P<.05, **P<.01, ***P<.001.



The logistic regression analyses presented in Table 2 demonstrate the predictive capacity of various diagnostic factors for frailty in older adults. The Nagelkerke R^2 values indicate the proportion of variance in frailty status explained by each model, both before and after adjusting for age and sex. The initial model for US shows that US measurements of the rectus femoris explained 10.1% of the variance in frailty (Nagelkerke R^2=0.101; P=.009). After adjusting for age and sex, the explained variance increases to 20.6% (adjusted Nagelkerke R^2=0.206; P=.002), indicating a stronger association when these factors are considered. WBPhA exhibited a significant association with frailty, explaining 16.4% of the variance in the unadjusted model (Nagelkerke R^2=0.164; P=.001). The adjusted model further enhanced the explained variance to 24.5% (adjusted Nagelkerke R^2=0.245; P<.001), underscoring its robustness as a predictive factor. Both left leg lean mass and right leg lean mass (LeanM RL) were significant predictors of frailty. The unadjusted models explained 12.4% (P=.003) and 16.4% (P<.001) of the variance, respectively. After adjustment, the explained variance increased to 21.5% for the left leg (P=.001) and 24.6% for the right leg (P<.001).

Table . Results of the logistic regression analyses (n=94).

	Nagelkerke R^2	*P* value	Adjusted Nagelkerke R^{2a}	*P* value
Ultrasound on rectus femoris	0.101	.009	0.206	.002
WBPhA[b]	0.164	.001	0.245	<.001
Lean mass, left leg	0.124	.003	0.215	.001
Lean mass, right leg	0.164	<.001	0.246	<.001

[a]Adjusted Nagelkerke R^2 included age as a covariable and sex as a factor.

[b]WBPhA: whole-body phase angle.

The ROC analyses indicated that all evaluated measures demonstrated a high degree of specificity in detecting frailty in older adults. It is notable that unadjusted muscle US demonstrated a high level of specificity, although sensitivity was relatively low. The WBPhA and LeanM RL also exhibited significant predictive capabilities, with AUC and *P* values indicating good discrimination. When adjusting for sex and age, there was a notable improvement in sensitivity, particularly in the case of LeanM RL, which maintained statistical significance (*P*=.045). These findings suggest that the lean mass of the right leg, adjusted for demographic factors, may be a promising indicator for predicting frailty in this population. On the other hand, it is worth noting that the adjusted WBPhA presented the highest Youden index value (0.417), which suggests that it may be the most effective measure in terms of balancing sensitivity and specificity for predicting frailty. Similarly, the adjusted LeanM RL also had a high Youden index (0.296), which establishes it as another promising measure.

Discussion

Principal Findings

This study provides significant evidence on the predictive capacity of body composition variables and their relationship with frailty syndrome in older adults. Technology tools such as BIA, WBPhA, US of the anterior rectus quadriceps muscle thickness, and DXA, were used in this study. The older adult frailty status was calculated using the PowerFrail mobile app [15] (Figure 1). This app is the first scientifically based app that allows the assessment of the muscle power and frailty of older adults in a simple way. Our results showed that these parameters are useful predictors for identifying frailty, in line with previous findings.

This study found several significant correlations between variables related to body composition, physical performance, and cognitive status in female participants, but these were not the same as the values observed in their male counterparts. This may be because women experience more functional limitations than men during the aging process [31]. Women tend to suffer a greater decline in physical function with age, which has been attributed to hormonal changes, particularly the decrease in estrogen levels during menopause [32]. Menopause is associated with changes in body composition characterized by an increase in body fat and a progressive decrease in muscle mass and strength [33]. These alterations can lead to a higher prevalence of sarcopenia and frailty in older women compared to men.

Based on these postmenopausal changes, muscle and fat composition might be more closely related to physical performance in women, which we observed in our correlation results from Figures 2-4. Specifically, the associations between body composition measures (evaluated through BIA, DXA, and US) and expressions of strength, such as lower limb power from the mobile app and handgrip strength, were more pronounced in women. This suggests that changes in muscle quality and quantity may have a greater impact on the physical function of older women.

Although previous research has indicated that muscle quality contributes to physical capacity in older adults, the effect of sex differences on this association has not been thoroughly investigated [34]. Some studies have suggested that muscle quality independently predicts physical function in older men but not in women [35]. However, our findings differ, as we observed significant correlations in women. Unlike previous studies, we included analyses of physical function status, which may explain the different results. The underlying mechanisms of these sex-specific differences still need to be investigated and clarified.

Contributing factors may include differences in muscle structure composition, hormonal influences, and neuromuscular activation patterns between men and women [36]. Additionally, the accuracy and sensitivity of measurement techniques, such as BIA and US, may vary between sexes due to differences in body fat distribution and hydration status [37]. Further research is needed to explore these factors and understand how they influence the relationship between body composition and physical function in older adults.

The connection between body composition and frailty indicators has been the focus of extensive research. However, to our knowledge, no studies have simultaneously compared DXA, BIA, and US, specifically for the identification of frailty syndrome in a primary care setting. Our study addresses this gap by evaluating the predictive capacities of these 3 modalities, offering a comprehensive analysis that can inform clinical practice regarding the most suitable and practical tools for frailty assessment in older adults. Although previous studies have compared two modalities, our inclusion of a third (US) provides a broader perspective on body composition assessment tools available for primary care settings. For instance, research has demonstrated that muscle US is an emerging tool for diagnosing sarcopenia, with studies summarizing its diagnostic accuracy [38]. Additionally, studies have evaluated the reliability and

validity of sarcopenia diagnosis using BIA compared with the gold standard, DXA, assessing the predictive accuracy of BIA for diagnosis [39,40]. However, these studies did not include US in their comparisons. Rossini-Venturini et al [41] highlighted that the anthropometric prediction equations developed in their study provide a reliable, practical, and low-cost instrument to assess the components that change the most during the aging process, corroborating our findings. This perspective emphasizes the significance of considering diverse elements of body composition in the evaluation of health among older adults. Although some studies have recognized muscle mass, assessed via DXA or BIA, as a critical element in forecasting frailty [42,43], our analysis demonstrated that lower extremity lean mass (left left lean mass and LeanM RL) did reveal a direct relationship in the models. The present results are in line with findings that indicate muscle mass reduction alone does not suffice to predict frailty without factoring in physical performance [44]. In fact, physical performance, which can be measured through functional tests such as gait speed or grip strength, is vital for evaluating frailty status among older adults [45].

There has been a growing acknowledgment in recent literature regarding the significance of assessing both the quantity and quality of muscle. Xu et al [46] demonstrated that body composition, encompassing both muscle mass and quality, is significantly associated with frailty in older adult inpatients. The current literature suggests that it is not enough to analyze the quantity of muscle mass in absolute terms; quality is equally important. Quality can be assessed through methods such as US and BIA, which provide insights into muscle integrity and performance in terms of its functional capacity. In this context, analyzing muscle quality through US could effectively complement DXA, which has traditionally been considered the gold standard in the assessment of body composition. Although DXA provides valuable information on lean mass and fat mass, it does not provide details on muscle distribution and quality, critical for understanding frailty in older adults [47].

Moreover, our findings indicated that muscle strength and gait speed are important indicators of frailty, corroborating the work of Tsukasaki et al [48], who found a strong association between muscle strength, gait speed, and cross-sectional muscle area determined by midthigh computed tomography. These findings reinforce the idea that comprehensive evaluations of muscle function should be integrated into frailty assessments. Therefore, muscle composition emerges as a potential public health assessment by enabling the clinical quantification of muscle mass and an estimation of physical function in the older adult population [49].

Our findings demonstrated that all measures exhibited high specificity but varying sensitivity in detecting frailty (Table 3). Our ROC analysis showed moderate predictive ability for WBPhA, US thickness, and lean mass of legs from DXA, with an AUC between 0.678 and 0.749. This aligns with the results from previous studies [50,51], which emphasized that models combining body composition measurements with physical performance tests can improve the predictive ability of frailty. Unadjusted US showed perfect specificity but low sensitivity (10.7%), indicating it is highly effective at correctly identifying nonfrail individuals but less capable of detecting those who are frail. This aligns with previous studies suggesting that muscle US, while precise in measuring muscle thickness, may have limitations in sensitivity due to operator dependency and variability [52].

Table . Results of receiver operating characteristic analysis (n=94)

	Sensitivity	Specificity	Youden index[a]	AUC[b]	*P* value
US[c]	0.107	1.000	0.107	0.678	.02
US (adjusted)[d]	0.321	0.939	0.260	0.732	.24
WBPhA[e]	0.240	0.952	0.192	0.704	.003
WBPhA (adjusted)[d]	0.480	0.937	0.417	0.762	.09
LeanM LL[f]	0.107	0.985	0.092	0.683	.009
LeanM LL (adjusted)[d]	0.250	0.924	0.174	0.741	.17
LeanM RL[g]	0.286	0.939	0.225	0.703	.003
LeanM RL (adjusted)[d]	0.357	0.939	0.296	0.749	.045

[a]Youden index = sensitivity + specificity – 1.

[b]AUC: area under the curve.

[c]US: ultrasound on rectus femoris.

[d]Model adjusted by sex and age.

[e]WBPhA: whole-body phase angle.

[f]LeanM LL: left leg lean mass.

[g]LeanM RL: right leg lean mass.

Adjusting the US measure for age and sex increased sensitivity to 32.1% but reduced specificity to 93.9%, and the adjusted model did not reach statistical significance (*P*=.24). This suggests that while adjustments improve sensitivity, they may

not be sufficient to make US a standalone diagnostic tool for frailty in primary care settings. The WBPhA showed an unadjusted sensitivity of 24% and specificity of 95.2%, with an AUC of 0.704 (*P*=.003). After adjusting for age and sex, sensitivity improved to 48%, specificity slightly decreased to 93.7%, and AUC increased to 0.762, although the *P* value was nonsignificant (*P*=.09). These results suggest that WBPhA, particularly when adjusted for sex and age factors, has potential as a screening tool for frailty. Previous research has shown that lower phase angle values are associated with decreased muscle function, poor nutritional status, and higher frailty risk [12,53]. The noninvasive nature and ease of use of BIA make WBPhA a practical option for primary care, although standardization of measurement protocols is necessary.

LeanM RL emerged as a significant predictor when adjusted for age and sex (*P*=.045), with sensitivity and specificity of 35.7% and 93.9%, respectively, and an AUC of 0.749. This indicates that the adjusted LeanM RL model had a good balance between sensitivity and specificity and may be valuable in predicting frailty. The importance of lower limb muscle mass in frailty assessment is well-documented. For instance, [54] reported that decreased appendicular lean mass is associated with physical disability and increased risk of adverse outcomes in older adults. However, the use of DXA in primary care is limited due to cost, accessibility, and exposure to low-dose radiation.

In comparison, the unadjusted left leg lean mass showed similar specificity (98.5%) but low sensitivity (10.7%), and the adjusted model did not achieve statistical significance. This suggests that while lean mass measurements are informative, the right leg may provide more predictive value than the left in this context, possibly due to dominance or functional differences, although further research is needed to confirm this observation. Overall, our results indicated that LeanM RL, adjusted for age and sex, may be the most effective measure among those studied for predicting frailty. This is significant because identifying reliable, accessible markers for frailty is crucial for early intervention. Given the limitations of DXA in primary care, exploring alternative methods to estimate lean mass, such as predictive equations or portable devices, could enhance feasibility. The high specificity observed across all measures suggests they are effective in ruling out frailty in individuals without frailty. However, the variable sensitivity underscores the need for multicomponent assessment tools. The Comprehensive Geriatric Assessment remains the gold standard for frailty evaluation but is resource-intensive [55]. Incorporating measures like WBPhA and simplified lean mass assessments could enhance screening efficiency in primary care.

Limitations, Clinical Implications, and Future Directions

Limitations of the study include the sample size and homogeneity in some variables, such as BMI and arm lean mass (left arm and right arm), which could have affected the accuracy of the predictive results. Heterogeneity in these measurements has been reported in previous research [56], suggesting that variability in body composition could influence frailty prediction. There is potential for the integration of advanced body composition analysis tools and mobile technology into primary care services for the early identification and management of frailty syndrome in older adults. Methods such as WBPhA, US of the rectus femoris muscle, and DXA analysis proved to be effective predictors of frailty, offering high specificity in detection. The implementation of the PowerFrail mobile app enables rapid and personalized assessments, optimizing preventive interventions. These tools not only improve the accuracy of clinical evaluations but also reduce costs associated with frailty-related complications, promoting healthier aging and alleviating the burden on public health care systems. In addition, while this study incorporated a range of assessment tools, it is essential for future research to concentrate on validating these measures in larger and more diverse cohorts to ensure their wider applicability. The cross-sectional design of this study further restricts the ability to infer causal relationships between body composition variables and frailty. Longitudinal investigations could provide valuable information on how variations in body composition affect the progression of frailty over time. Understanding these dynamics could facilitate developing more effective and personalized interventions for the older adult population. Finally, future studies should explore the integration of these mHealth technologies with advanced body composition analysis systems to optimize early detection and management of frailty.

Conclusions

The findings of this study reinforce the utility of various body composition evaluations, such as WBPhA, LeanM RL measured by DXA, and quadriceps thickness assessed by US, as effective indicators for predicting frailty in older adults, aligning with previous research. However, our results highlight the necessity of not relying exclusively on muscle mass as a predictor of frailty. It is essential to incorporate assessments of muscle function and physical performance into clinical evaluations to enhance the accuracy of identifying individuals susceptible to frailty. The combination of tools such as WBPhA, bioimpedance, and US, along with the PowerFrail app, could provide a more complete and accurate assessment of the health status of older adults, allowing effective preventive interventions to be implemented in primary care services.

Acknowledgments

This study received funding from the Call for Biomedical Research and Innovation Projects from the Biomedical Research Institute of Cádiz, Andalusian Health System Grant (PP11-004-2023); PUENTE Project Call for the Enhancement of Proposals from the State Plan of the University (PR2022-018); and Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI 2020; DOC_01412) from the Andalusian Government.

Conflicts of Interest

None declared.

References

1. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. J Gerontol A Biol Sci Med Sci 2006 Mar;61(3):262-266. [doi: 10.1093/gerona/61.3.262] [Medline: 16567375]
2. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
3. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
4. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine defined by frailty. Clin Geriatr Med 2011 Feb;27(1):17-26. [doi: 10.1016/j.cger.2010.08.008] [Medline: 21093719]
5. Galán-Mercant A, Cuesta-Vargas AI. Detección precoz de la fragilidad, tecnología aplicada al movimiento humano para la prevención de la discapacidad [Article in Spanish]. Fisioterapia 2017 May;39(3):135-136. [doi: 10.1016/j.ft.2016.10.002]
6. Puts MTE, Toubasi S, Andrew MK, et al. Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a scoping review of the literature and international policies. Age Ageing 2017 May 1;46(3):383-392. [doi: 10.1093/ageing/afw247] [Medline: 28064173]
7. Dent E, Morley JE, Cruz-Jentoft AJ, et al. International Clinical Practice Guidelines for Sarcopenia (ICFSR): screening, diagnosis and management. J Nutr Health Aging 2018;22(10):1148-1161. [doi: 10.1007/s12603-018-1139-9] [Medline: 30498820]
8. Rolland Y, Czerwinski S, Abellan Van Kan G, et al. Sarcopenia: its assessment, etiology, pathogenesis, consequences and future perspectives. J Nutr Health Aging 2008;12(7):433-450. [doi: 10.1007/BF02982704] [Medline: 18615225]
9. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis. Age Ageing 2010 Jul 1;39(4):412-423. [doi: 10.1093/ageing/afq034]
10. Beaudart C, McCloskey E, Bruyère O, et al. Sarcopenia in daily practice: assessment and management. BMC Geriatr 2016 Oct 5;16(1):170. [doi: 10.1186/s12877-016-0349-4] [Medline: 27716195]
11. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 Feb;27(1):5-15. [doi: 10.1016/j.clnu.2007.10.007] [Medline: 18061312]
12. Matias CN, Nunes CL, Francisco S, et al. Phase angle predicts physical function in older adults. Arch Gerontol Geriatr 2020;90:104151. [doi: 10.1016/j.archger.2020.104151] [Medline: 32563736]
13. Gupta M, Lehl SS, Lamba AS. Ultrasonography for assessment of sarcopenia: a primer. J Midlife Health 2022;13(4):269-277. [doi: 10.4103/jmh.jmh_234_22] [Medline: 37324795]
14. Buckinx F, Landi F, Cesari M, et al. Pitfalls in the measurement of muscle mass: a need for a reference standard. J Cachexia Sarcopenia Muscle 2018 Apr;9(2):269-278. [doi: 10.1002/jcsm.12268] [Medline: 29349935]
15. Alcazar J, Losa-Reyna J, Rodriguez-Lopez C, et al. The sit-to-stand muscle power test: an easy, inexpensive and portable procedure to assess muscle power in older people. Exp Gerontol 2018 Oct 2;112:38-43. [doi: 10.1016/j.exger.2018.08.006] [Medline: 30179662]
16. Losa-Reyna J, Alcazar J, Rodríguez-Gómez I, et al. Low relative mechanical power in older adults: an operational definition and algorithm for its application in the clinical setting. Exp Gerontol 2020 Dec;142:111141. [doi: 10.1016/j.exger.2020.111141] [Medline: 33127413]
17. Velazquez-Diaz D, Arco JE, Ortiz A, et al. Use of artificial intelligence in the identification and diagnosis of frailty syndrome in older adults: scoping review. J Med Internet Res 2023 Oct 20;25:e47346. [doi: 10.2196/47346] [Medline: 37862082]
18. Liu N, Yin J, Tan SSL, Ngiam KY, Teo HH. Mobile health applications for older adults: a systematic review of interface and persuasive feature design. J Am Med Inform Assoc 2021 Oct 12;28(11):2483-2501. [doi: 10.1093/jamia/ocab151] [Medline: 34472601]
19. Soto-Bagaria L, Eis S, Pérez LM, et al. Mobile applications to prescribe physical exercise in frail older adults: review of the available tools in app stores. Age Ageing 2023 Dec 1;52(12):afad227. [doi: 10.1093/ageing/afad227] [Medline: 38157286]
20. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
21. Bohannon RW. Sit-to-stand test for measuring performance of lower extremity muscles. Percept Mot Skills 1995 Feb;80(1):163-166. [doi: 10.2466/pms.1995.80.1.163] [Medline: 7624188]
22. Losa-Reyna J, Baltasar-Fernandez I, Alcazar J, et al. Effect of a short multicomponent exercise intervention focused on muscle power in frail and pre frail elderly: a pilot trial. Exp Gerontol 2019 Jan;115:114-121. [doi: 10.1016/j.exger.2018.11.022] [Medline: 30528641]
23. Mateos-Angulo A, Galán-Mercant A, Cuesta-Vargas AI. Muscle thickness contribution to sit-to-stand ability in institutionalized older adults. Aging Clin Exp Res 2020 Aug;32(8):1477-1483. [doi: 10.1007/s40520-019-01328-x] [Medline: 31463929]
24. Villarejo A, Puertas-Martín V. Utilidad de los test breves en el cribado de demencia [Article in Spanish]. Neurología 2011 Sep;26(7):425-433. [doi: 10.1016/j.nrl.2010.12.002]

25. Muñoz Díaz B, Molina-Recio G, Romero-Saldaña M, et al. Validation (in Spanish) of the Mini Nutritional Assessment survey to assess the nutritional status of patients over 65 years of age. Fam Pract 2019 Mar 20;36(2):172-178. [doi: 10.1093/fampra/cmy051] [Medline: 29873713]
26. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
27. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
28. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011 Jul;40(4):423-429. [doi: 10.1093/ageing/afr051] [Medline: 21624928]
29. Martínez-Camblor P, Pardo-Fernández JC. The Youden index in the generalized receiver operating characteristic curve context. Int J Biostat 2019 Apr 3;15(1). [doi: 10.1515/ijb-2018-0060] [Medline: 30943172]
30. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988 Jun 3;240(4857):1285-1293. [doi: 10.1126/science.3287615] [Medline: 3287615]
31. Iannuzzi-Sucich M, Prestwood KM, Kenny AM. Prevalence of sarcopenia and predictors of skeletal muscle mass in healthy, older men and women. J Gerontol A Biol Sci Med Sci 2002 Dec;57(12):M772-M777. [doi: 10.1093/gerona/57.12.m772] [Medline: 12456735]
32. Maltais ML, Desroches J, Dionne IJ. Changes in muscle mass and strength after menopause. J Musculoskelet Neuronal Interact 2009;9(4):186-197. [Medline: 19949277]
33. Sipilä S, Narici M, Kjaer M, et al. Sex hormones and skeletal muscle weakness. Biogerontology 2013 Jun;14(3):231-245. [doi: 10.1007/s10522-013-9425-8] [Medline: 23636830]
34. Barbat-Artigas S, Rolland Y, Vellas B, Aubertin-Leheudre M. Muscle quantity is not synonymous with muscle quality. J Am Med Dir Assoc 2013 Nov;14(11):852. [doi: 10.1016/j.jamda.2013.06.003] [Medline: 23896368]
35. Goodpaster BH, Park SW, Harris TB, et al. The loss of skeletal muscle strength, mass, and quality in older adults: the health, aging and body composition study. J Gerontol A Biol Sci Med Sci 2006 Oct;61(10):1059-1064. [doi: 10.1093/gerona/61.10.1059] [Medline: 17077199]
36. Clark BC, Manini TM. Sarcopenia =/= dynapenia. J Gerontol A Biol Sci Med Sci 2008 Aug;63(8):829-834. [doi: 10.1093/gerona/63.8.829] [Medline: 18772470]
37. Kyle UG, Genton L, Slosman DO, Pichard C. Fat-free and fat mass percentiles in 5225 healthy subjects aged 15 to 98 years. Nutrition 2001;17(7-8):534-541. [doi: 10.1016/s0899-9007(01)00555-x] [Medline: 11448570]
38. Fu H, Wang L, Zhang W, Lu J, Yang M. Diagnostic test accuracy of ultrasound for sarcopenia diagnosis: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2023 Feb;14(1):57-70. [doi: 10.1002/jcsm.13149] [Medline: 36513380]
39. Lee JH, Kim HJ, Han S, Park SJ, Sim M, Lee KH. Reliability and agreement assessment of sarcopenia diagnosis through comparison of bioelectrical impedance analysis and dual-energy X-ray absorptiometry. Diagnostics (Basel) 2024 Apr 25;14(9):899. [doi: 10.3390/diagnostics14090899] [Medline: 38732314]
40. Lee SY, Ahn S, Kim YJ, et al. Comparison between dual-energy X-ray absorptiometry and bioelectrical impedance analyses for accuracy in measuring whole body muscle mass and appendicular skeletal muscle mass. Nutrients 2018 Jun 7;10(6):738. [doi: 10.3390/nu10060738] [Medline: 29880741]
41. Rossini-Venturini AC, Veras L, Abdalla PP, et al. Multicompartment body composition analysis in older adults: a cross-sectional study. BMC Geriatr 2023 Feb 9;23(1):87. [doi: 10.1186/s12877-023-03752-1] [Medline: 36759773]
42. Kołodziej M, Sebastjan A, Ignasiak Z. Appendicular skeletal muscle mass and quality estimated by bioelectrical impedance analysis in the assessment of frailty syndrome risk in older individuals. Aging Clin Exp Res 2022 Sep;34(9):2081-2088. [doi: 10.1007/s40520-021-01879-y] [Medline: 34118025]
43. Kołodziej M, Kozieł S, Ignasiak Z. The use of the bioelectrical impedance phase angle to assess the risk of sarcopenia in people aged 50 and above in Poland. Int J Environ Res Public Health 2022 Apr 13;19(8):4687. [doi: 10.3390/ijerph19084687] [Medline: 35457554]
44. Taylor JA, Greenhaff PL, Bartlett DB, Jackson TA, Duggal NA, Lord JM. Multisystem physiological perspective of human frailty and its modulation by physical activity. Physiol Rev 2023 Apr 1;103(2):1137-1191. [doi: 10.1152/physrev.00037.2021] [Medline: 36239451]
45. Jiménez-García JD, Ortega-Gómez S, Martínez-Amat A, Álvarez-Salvago F. Associations of balance, strength, and gait speed with cognitive function in older individuals over 60 years: a cross-sectional study. Appl Sci (Basel) 2024;14(4):1500. [doi: 10.3390/app14041500]
46. Xu L, Zhang J, Shen S, et al. Association between body composition and frailty in elder inpatients. Clin Interv Aging 2020;15:313-320. [doi: 10.2147/CIA.S243211] [Medline: 32184580]
47. Oba H, Matsui Y, Arai H, et al. Evaluation of muscle quality and quantity for the assessment of sarcopenia using mid-thigh computed tomography: a cohort study. BMC Geriatr 2021 Apr 13;21(1):239. [doi: 10.1186/s12877-021-02187-w] [Medline: 33849469]

48. Tsukasaki K, Matsui Y, Arai H, et al. Association of muscle strength and gait speed with cross-sectional muscle area determined by mid-thigh computed tomography - a comparison with skeletal muscle mass measured by dual-energy X-ray absorptiometry. J Frailty Aging 2020;9(2):82-89. [doi: 10.14283/jfa.2020.16] [Medline: 32259181]
49. Alarcón-Rivera M, Cornejo-Mella C, Cáceres-Aravena C, Concha-Cisternas Y, Fernández-Valero P, Guzmán-Muñoz E. Relationship between appendicular muscular mass index and physical function in older people. AIMS Public Health 2024;11(1):130-140. [doi: 10.3934/publichealth.2024006] [Medline: 38617413]
50. Masanés F, Rojano I Luque X, Salvà A, et al. Cut-off points for muscle mass - not grip strength or gait speed - determine variations in sarcopenia prevalence. J Nutr Health Aging 2017;21(7):825-829. [doi: 10.1007/s12603-016-0844-5] [Medline: 28717813]
51. Bahat G, Tufan A, Tufan F, et al. Cut-off points to identify sarcopenia according to European Working Group on Sarcopenia in Older People (EWGSOP) definition. Clin Nutr 2016 Dec;35(6):1557-1563. [doi: 10.1016/j.clnu.2016.02.002] [Medline: 26922142]
52. Pillen S, Tak RO, Zwarts MJ, et al. Skeletal muscle ultrasound: correlation between fibrous tissue and echo intensity. Ultrasound Med Biol 2009 Mar;35(3):443-446. [doi: 10.1016/j.ultrasmedbio.2008.09.016] [Medline: 19081667]
53. Norman K, Stobäus N, Pirlich M, Bosy-Westphal A. Bioelectrical phase angle and impedance vector analysis--clinical relevance and applicability of impedance parameters. Clin Nutr 2012 Dec;31(6):854-861. [doi: 10.1016/j.clnu.2012.05.008] [Medline: 22698802]
54. Studenski SA, Peters KW, Alley DE, et al. The FNIH sarcopenia project: rationale, study description, conference recommendations, and final estimates. J Gerontol A Biol Sci Med Sci 2014 May;69(5):547-558. [doi: 10.1093/gerona/glu010] [Medline: 24737557]
55. Ellis G, Gardner M, Tsiachristas A, et al. Comprehensive geriatric assessment for older adults admitted to hospital. Cochrane Database Syst Rev 2017 Sep 12;9(9). [doi: 10.1002/14651858.CD006211.pub3] [Medline: 28898390]
56. Cesari M, Gambassi G, van Kan GA, Vellas B. The frailty phenotype and the frailty index: different instruments for different purposes. Age Ageing 2014 Jan;43(1):10-12. [doi: 10.1093/ageing/aft160] [Medline: 24132852]

Abbreviations

AUC: area under the curve
BIA: bioelectrical impedance analysis
DXA: dual-energy X-ray absorptiometry
GS4m: gait speed over 4 meters at a normal pace
LeanM RL: right leg lean mass
mHealth: mobile health
MNA: Mini Nutritional Assessment
ROC: receiver operating characteristic
RPOW: relative muscle power
SPPB: Short Physical Performance Battery
US: ultrasound
WBPhA: whole-body phase angle

Edited by M Antoniou; submitted 29.10.24; peer-reviewed by C Garcia-Munoz, EO GaitÁn, MN Shalaby; revised version received 08.01.25; accepted 10.02.25; published 15.05.25.

Please cite as:
Ortiz-Navarro B, Losa-Reyna J, Mihaiescu-Ion V, Garcia-Romero J, Carrillo de Albornoz-Gil M, Galán-Mercant A
Identification of Target Body Composition Parameters by Dual-Energy X-Ray Absorptiometry, Bioelectrical Impedance, and Ultrasonography to Detect Older Adults With Frailty and Prefrailty Status Using a Mobile App in Primary Care Services: Descriptive Cross-Sectional Study
JMIR Aging 2025;8:e67982
URL: https://aging.jmir.org/2025/1/e67982
doi:10.2196/67982

Using Deep Learning to Perform Automatic Quantitative Measurement of Masseter and Tongue Muscles in Persons With Dementia: Cross-Sectional Study

Mahdi Imani[1], PhD, Biomed Eng; Miguel G Borda[2,3], MD, PhD; Sara Vogrin[1], MD, MSc; Erik Meijering[4], PhD; Dag Aarsland[2,5], MD, PhD; Gustavo Duque[6], MD, PhD

[1]Department of Medicine, Melbourne Medical School, University of Melbourne, St. Albans, Australia

[2]Centre for Age-Related Medicine (SESAM), Stavanger University Hospital, Stavanger, Norway

[3]Department of Neurology, Clínica Universidad de Navarra, Pamplona, Spain

[4]School of Computer Science and Engineering, UNSW Sydney, Sydney, Australia

[5]Department of Old Age Psychiatry, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, United Kingdom

[6]Bone, Muscle & Geroscience Group, Research Institute of the McGill University Health Centre, McGill University, 1001 Decarie Blvd, Room EM1.3226, Montreal, QC, Canada

Corresponding Author:
Gustavo Duque, MD, PhD
Bone, Muscle & Geroscience Group, Research Institute of the McGill University Health Centre, McGill University, 1001 Decarie Blvd, Room EM1.3226, Montreal, QC, Canada

Abstract

Background: Sarcopenia (loss of muscle mass and strength) increases adverse outcomes risk and contributes to cognitive decline in older adults. Accurate methods to quantify muscle mass and predict adverse outcomes, particularly in older persons with dementia, are still lacking.

Objective: This study's main objective was to assess the feasibility of using deep learning techniques for segmentation and quantification of musculoskeletal tissues in magnetic resonance imaging (MRI) scans of the head in patients with neurocognitive disorders. This study aimed to pave the way for using automated techniques for opportunistic detection of sarcopenia in patients with neurocognitive disorder.

Methods: In a cross-sectional analysis of 53 participants, we used 7 U-Net-like deep learning models to segment 5 different tissues in head MRI images and used the Dice similarity coefficient and average symmetric surface distance as main assessment techniques to compare results. We also analyzed the relationship between BMI and muscle and fat volumes.

Results: Our framework accurately quantified masseter and subcutaneous fat on the left and right sides of the head and tongue muscle (mean Dice similarity coefficient 92.4%). A significant correlation exists between the area and volume of tongue muscle, left masseter muscle, and BMI.

Conclusions: Our study demonstrates the successful application of a deep learning model to quantify muscle volumes in head MRI in patients with neurocognitive disorders. This is a promising first step toward clinically applicable artificial intelligence and deep learning methods for estimating masseter and tongue muscle and predicting adverse outcomes in this population.

(JMIR Aging 2025;8:e63686) doi:10.2196/63686

KEYWORDS

artificial intelligence; machine learning; sarcopenia; dementia; masseter muscle; tongue muscle; deep learning; head; tongue; face; magnetic resonance imaging; MRI; image; imaging; muscle; muscles; neural network; aging; gerontology; older adults; geriatrics; older adult health

Introduction

Age-related muscle wasting and neurodegeneration, clinically presented as sarcopenia and dementia, respectively, are the major drivers of frailty, falls, and disability in older adults worldwide [1]. Sarcopenia is characterized by loss of muscle mass, strength, and function in older adults. Aging is the leading risk factor, but conditions such as chronic diseases, inflammation, sedentarism, and malnutrition promote sarcopenia onset and progression [2]. Sarcopenia has a 10% overall prevalence globally in older persons, 29% in the community, 14%-33% in long-term care settings, and up to 50% in the very old (>80) [3,4]. Despite being a common and relevant health-related condition, it is unseen and underdiagnosed,

particularly in older persons with cognitive disorders. To diagnose sarcopenia, measurement of muscle mass, muscle performance, and strength is necessary [2]. Estimating muscle performance and strength is accessible and cheap with traditional methods, such as gait speed and grip strength, respectively [5]. However, techniques such as dual X-ray absorptiometry or body magnetic resonance imaging (MRI) are necessary to accurately assess lean or muscle mass. These methods can increase costs and time and are impractical in settings such as dementia clinics [6].

Dementia patients are highly affected by sarcopenia, with a prevalence of around 60% - 70% [7,8]. People living with neurodegenerative diseases are more prone to experience difficulties due to malnutrition, being sedentary, and falling; therefore, having sarcopenia increases the risk of adverse outcomes. Sarcopenia is not only a risk factor for adverse outcomes for those with dementia but also promotes cognitive loss in healthy older adults [9]. Therefore, diagnosing sarcopenia in people with neurodegenerative diseases is relevant and necessary.

Head MRI is a widely used diagnostic method for assessing dementia and Alzheimer disease (AD), as it offers intricate representations of the brain's anatomy and physiology. In clinical practice, MRI is often combined with other imaging techniques and cognitive assessments to support the diagnosis of these conditions. The utilization of MRI has seen an upward trend in recent times, as it has become an instrumental tool for the early detection and tracking of the evolution of dementia and AD. According to estimates, 60% - 80% of patients diagnosed with dementia or AD undergo MRI as part of their diagnostic evaluation [6].

Mastication and deglutition muscles such as the masseter and tongue are visible in brain MRI scans [10,11]. These muscles can reflect not only age-associated general muscle decline but also systemic processes due to highly complex interactions with the immune system and the inflammatory response, the nervous system, and the crossroads of several components of the frailty syndrome [12,13]. Indeed, in a previous publication, we have reported that manually segmented masseter predicts mortality and clinical short-term and long-term outcomes in several contexts [14]. Moreover, head muscles such as the tongue and masseter could be cost-effective alternatives to estimate muscle mass in dementia and other common conditions such as head cancer, stroke, or cranioencephalic trauma [14,15]. However, manual and semi-automatic techniques are labor-intensive and time-consuming, making the image processing task for large studies difficult, expensive, and, most importantly, impractical to apply in a clinical setting. Therefore, in the present study, we aimed to use MRI scans of the head opportunistically to develop an automated deep learning method to evaluate sarcopenia.

Methods

Population and Data Source

The "Dementia Study of Western Norway" (DemVest) is a long-term study between 2005 and 2013, with ongoing follow-up assessments. Participants were referred from dementia clinics in Hordaland and Rogaland and were insured by the same national insurance scheme. The study's methodology is described elsewhere [15]. Those with moderate or severe dementia, delirium, past bipolar or psychotic conditions, terminal illness, or newly diagnosed somatic diseases impacting cognition, function, or participation were excluded.

For this study, subjects with dementia with Lewy bodies (DLB) or mild AD who had baseline MRI scans were included. Out of 111 participants (85 AD and 26 DLB), 33 AD and 20 DLB participants with MRI images for brain and muscle measurement were selected based on the quality of the images and clear delineation of the regions of interest. The diagnosis of dementia was made in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria, and patients were classified as AD or DLB [16]. A mini-mental state examination score of ≥20 or a clinical dementia rating global score of 1 was chosen as the definition for mild dementia. The diagnosis was based on inclusion but could be modified with clinical evolution, consensus, and autopsy [15]. Participants were evaluated through structured assessments, and medical records were used to gather complete medical history and comorbidity data. In total, 56 participants had pathological diagnoses with 80% accuracy compared to clinical criteria, which reflects a reliable initial clinical diagnosis [17].

Ethical Considerations

This study was approved by the regional ethics committee (approval code: 2010/633) and the Norwegian authorities for the collection of medical data. All data were handled and kept following national data privacy protocols. All participants signed an informed consent form before inclusion in the study.

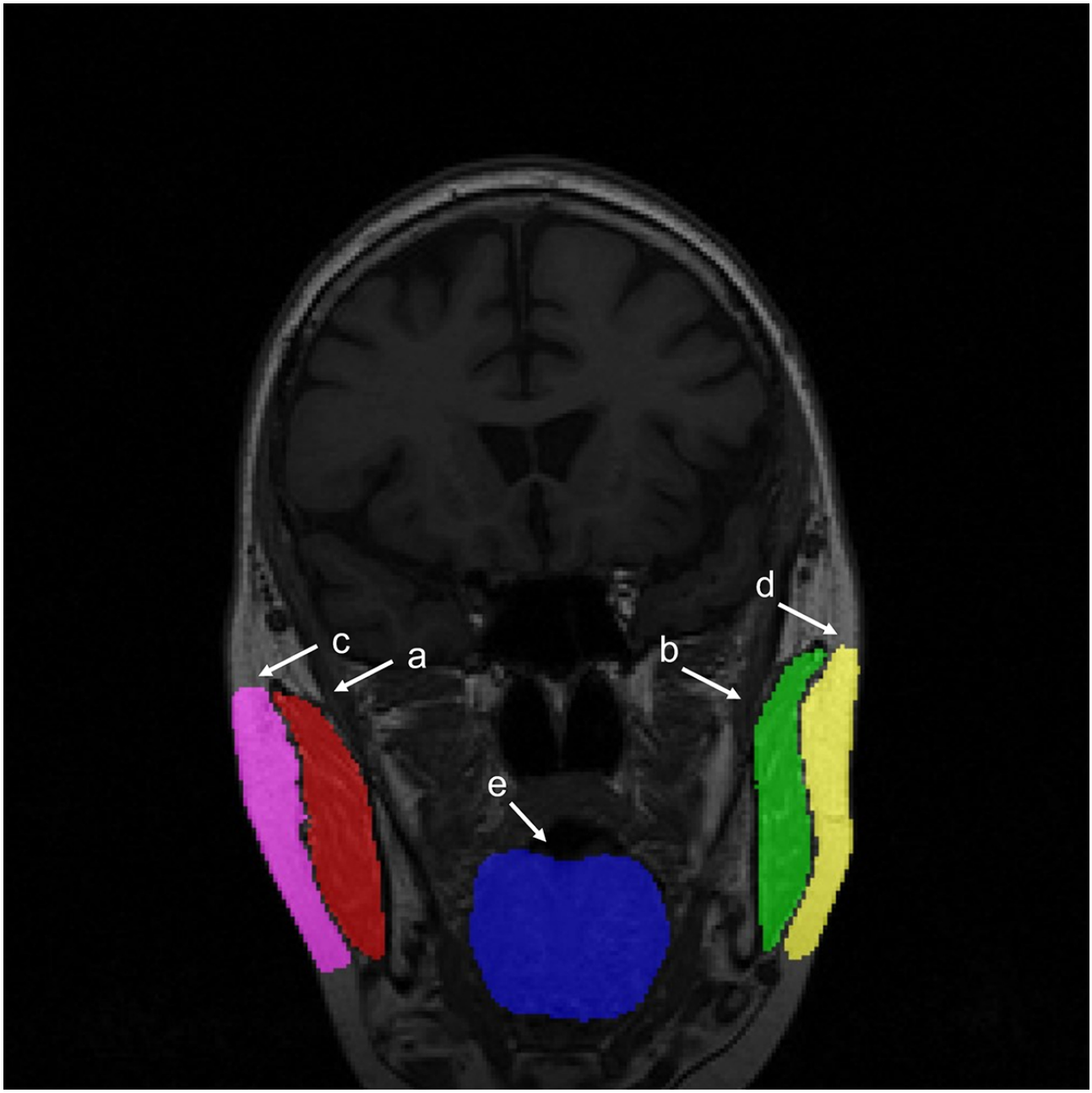
Imaging

All images were acquired at baseline. A 1.5-T Philips Intera-scanner was used to obtain MRI images. The acquisition protocol for 3D T1-weighted sequence was as follows: flip angle of 30°, repetition time/echo time of 10.0/4.6 ms, number of excitations of 2, 2-mm slice thickness with 1-mm spacing between the slices (1-mm slices with no gap), matrix of 256×256 pixels, and field of view of 26 cm. Those with movement artifacts and inadequate image quality were removed from the data using visual quality checks. A standardized preprocessing method for harmonizing multiple collections of MRIs was applied, which consisted of movement correction and intensity normalization following previously validated techniques [15].

Ground Truth Image Segmentation

Ground truth (GT) images were segmented using interactive pixel techniques made available by SliceOmatic software (TomoVision) following a manual method previously reported [18]. The size of the masseter muscle was used as a reference for the selection criteria of the slices. In total, 5 slices were selected from the ones with both right and left masseter muscles at their largest. For each slice, 5 tissues were segmented: left and right masseter muscles, left and right subcutaneous fat, and tongue muscle (Figure 1). The masseter muscle on each side was used as a reference to segment subcutaneous fat.

Figure 1. Example of segmented tissues overlaid on the original MRI. (**A**) Right masseter muscle, (**B**) left masseter muscle, (**C**) right subcutaneous fat, (**D**) left subcutaneous fat, and (**E**) tongue muscle.



Network Architecture

We studied and compared 6 different U-Net-like architectures. The original U-Net architecture [19] was first designed to segment medical images, and many other researchers have tried to improve its performance by integrating additional techniques into its architecture [20]. U-Net consists of a contracting path (encoder) and an expansive path (decoder) with skip connections between these 2 paths. The network learns features from the provided image and the mask at increasingly higher spatial scales by gradually down-sampling to lower resolutions through the encoding path. The expansive path then gradually increases the resolution of the output from the encoding path to the original image size, resulting in a probability map as an output, indicating the chance of each pixel belonging to a specific tissue. One important feature of U-Net is its skip connections, which concatenate feature maps from the encoding path to the corresponding block in the expansive path, making it possible to maintain small details crucial in medical image segmentation.

We have included 5 variants of U-Net in this study, including Attention U-Net, Dense U-Net, Residual U-Net, Inception U-Net, and U-Net++. Attention U-Net is desirable since it allows the model to focus on specific objects and ignore unnecessary areas [21]. In Dense U-Net, the traditional U-Net blocks are replaced with a dense block, enabling the model to segment objects with greater distinction. This feature is important in medical practice since tissues are often very close and sometimes overlap [22]. Residual U-Net architecture tries to tackle the vanishing gradient issue, a common problem in designing deep neural networks [23]. In most cases, the same organ's size can vary between patients, which can cause limitations on the segmentation capability of the model. By using filters with

different sizes, Inception U-Net attempts to overcome this problem [24]. Lastly, U-Net++ aids the classic U-Net model to more accurately segment images by providing semantic information from a dense network of skip connections as an intermediary grid between the encoding and decoding paths [25]. We also included a Wide U-Net architecture to eliminate the effect of increased trainable parameters. This model has the same architecture as U-Net but with more feature maps per layer (30, 60, 120, 240, and 480) compared to the original U-Net (16, 32, 64, 128, and 256). Hence, it will serve as a control to compare the models with larger numbers of trainable parameters with the base U-Net (Table 1).

Table . Number of trainable parameters for each model.

Model	Trainable parameters
U-Net	2,164,390
Attention U-Net	2,233,270
U-Net++	2,555,702
Inception U-Net	5,529,526
Residual U-Net	6,877,110
Dense U-Net	7,666,320
Wide U-Net	7,596,306

Training Procedure

All models were trained with the mini-batch stochastic gradient descent algorithm using the Adam optimizer. A batch size of 8 was selected. The learning rate was 0.0001. The training was done for 200 epochs. The values of hyperparameters were empirically tuned for best performance. Categorical cross-entropy (CSE) was selected as the loss function for this study [26]. The CSE loss function minimizes the distance between 2 distributions (the predicted labels and the GT labels). CSE is one of the most popular loss functions for image segmentation and has shown excellent performance in muscle segmentation [27,28]. All experiments were implemented with open-source software: Python (version 3.7.13), TensorFlow (version 2.8.2), and Keras (version 2.8.0).

Model Evaluation

The results of the experiments were evaluated using 2 main measures: Dice similarity coefficient (DSC) and average symmetric surface distance (ASSD). The DSC represents the agreement between the GT labels and predicted labels that models generate:

$$DSC_{P,G} = 2(P \cap G)(P+G)$$

where $\cap$ is the intersection and P and G are the 2 labels. DSC ranges between 0 and 1, where 0 indicates no agreement and 1 indicates perfect agreement. In our study, DSC is presented as a percentage.

The ASSD measures the average distance from pixels on the boundary of predicted labels to corresponding pixels on the boundary of the GT labels:

where BP and BG are the boundaries of predicted labels and corresponding reference labels, respectively. d(v,B) is the shortest Euclidean distance between voxel v and boundary B. An ASSD of 0 indicates a perfect match between predicted and reference labels. The ASSD was measured in mm.

We used k-fold cross-validation in evaluating the models. This technique splits the data set into k subsets (folds). The deep learning models are trained on all but one of the subsets (k–1), and then the models are evaluated on the subset that was not used for training. This process is repeated k times, and the average of the results is reported. We used k=10.

Additional metrics, including Jaccard coefficient, precision, recall, sensitivity, specificity, and F_1-score, are presented in Multimedia Appendix 1. The results are shown as mean and standard deviation and median and interquartile range across k-fold (k=10) cross-validation.

Statistical Methods

We explored the association between the BMI and the muscle and fat areas and volumes in the MRI images using individual linear regression models adjusted by sex and age. No adjustments for multiple testing were made. All assumptions were checked. All *P*-values were evaluated at a 5% level. The analysis and graphs were carried out using R (version 4.2.2).

Results

Segmentation

DSC and ASSD were used to quantitatively analyze the segmentation results (Figure 2 and Table 2). The Dense U-Net model has a higher average DSC and lower ASSD than other models. In contrast, Attention U-Net has the lowest average DSC score compared to other models. This result is confirmed by the higher ASSD for Attention U-Net for all tissues. Other models have shown almost similar results for all areas. Additional metrics have been presented in Multimedia Appendix 1. The results from these metrics confirm the findings of the study mentioned in this section; hence, they were omitted from the main manuscript.

Figure 2. Box plot of k-fold (k=10) cross-validation results for attention U-Net, dense U-Net, inception U-Net, residual U-Net, U-Net, U-Net++, and wide U-Net. Top: DSC in percentage. Bottom: ASSD in mm.

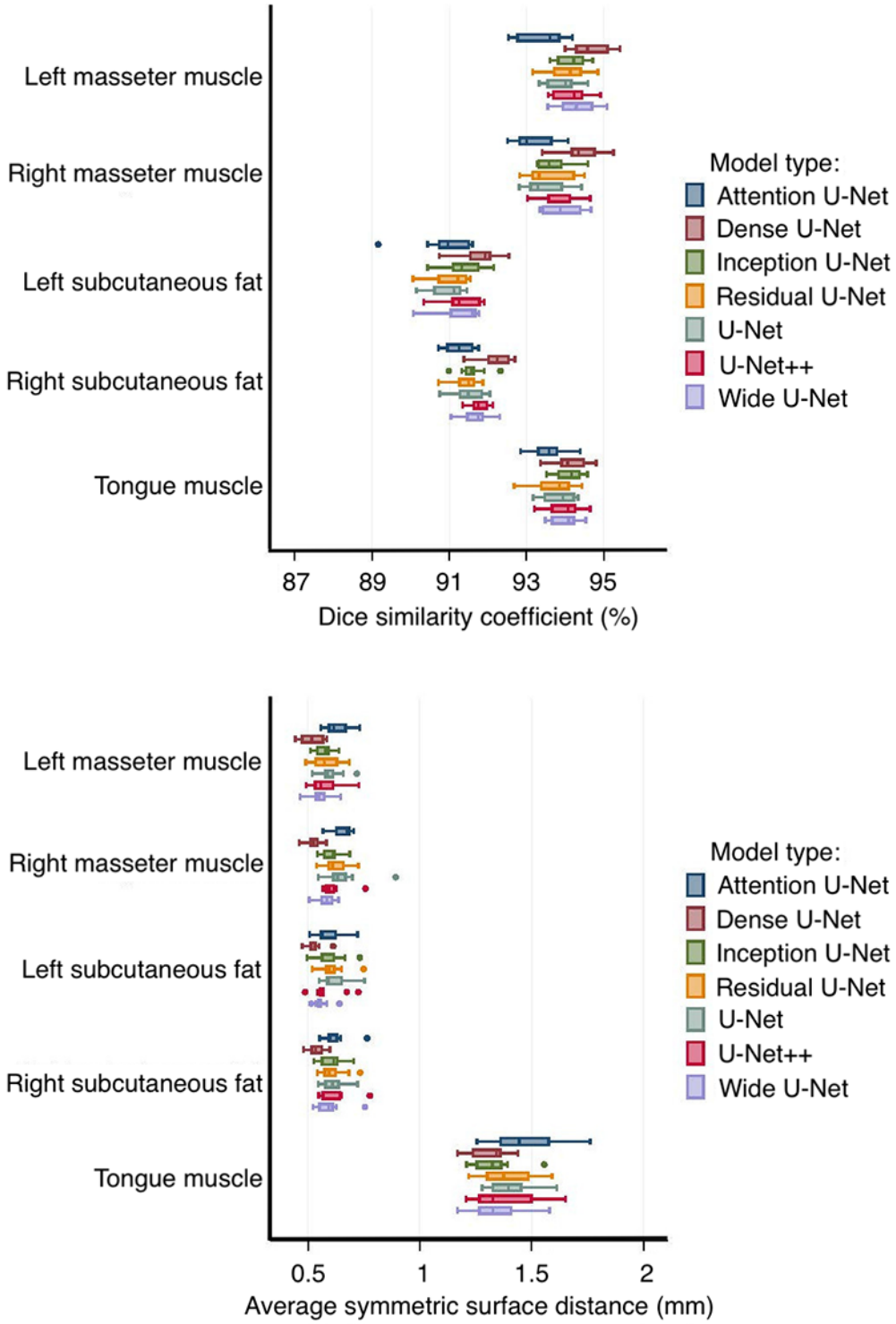


Table . Mean DSC and ASSD for test and validation results for k-fold (k=10) cross-validation. Standard deviation for the measurements in this table is presented in Multimedia Appendix 1.

	Model	Test set					Validation set				
		T[a]	LM[b]	RM[c]	LSF[d]	RSF[e]	T	LM	RM	LSF	RSF
DSC[f] (%)	Attention U-Net	93.57	93.4	93.19	90.9	91.26	93.43	93.38	93.23	91.05	91.43
	Dense U-Net	94.12	94.66	94.4	91.76	92.22	94.07	94.7	94.49	91.89	92.42
	Inception U-Net	94.07	94.12	93.69	91.36	91.59	93.85	94.27	93.73	91.45	91.81
	Residual U-Net	93.71	94.03	93.53	91.05	91.46	93.74	94.1	93.67	91.35	91.68
	U-Net	93.83	93.94	93.46	90.97	91.49	93.88	93.93	93.52	91.18	91.67
	Wide U-Net	94	94.29	93.93	91.35	91.68	94.03	94.3	93.99	91.59	92.01
ASSD[g] (mm)	Attention U-Net	1.47	0.63	0.66	0.59	0.62	1.53	0.67	0.65	0.6	0.59
	Dense U-Net	1.3	0.52	0.53	0.53	0.54	1.37	0.51	0.53	0.52	0.54
	Inception U-Net	1.33	0.57	0.61	0.59	0.61	1.44	0.56	0.63	0.6	0.59
	Residual U-Net	1.39	0.58	0.62	0.61	0.61	1.38	0.58	0.63	0.58	0.6
	U-Net	1.41	0.6	0.65	0.62	0.61	1.45	0.6	0.64	0.61	0.59
	Wide U-Net	1.35	0.55	0.58	0.56	0.6	1.34	0.55	0.58	0.56	0.57

[a]T: tongue muscle.

[b]LM: left masseter muscle.

[c]RM: right masseter muscle.

[d]LSF: left subcutaneous fat.

[e]RSF: right subcutaneous fat.

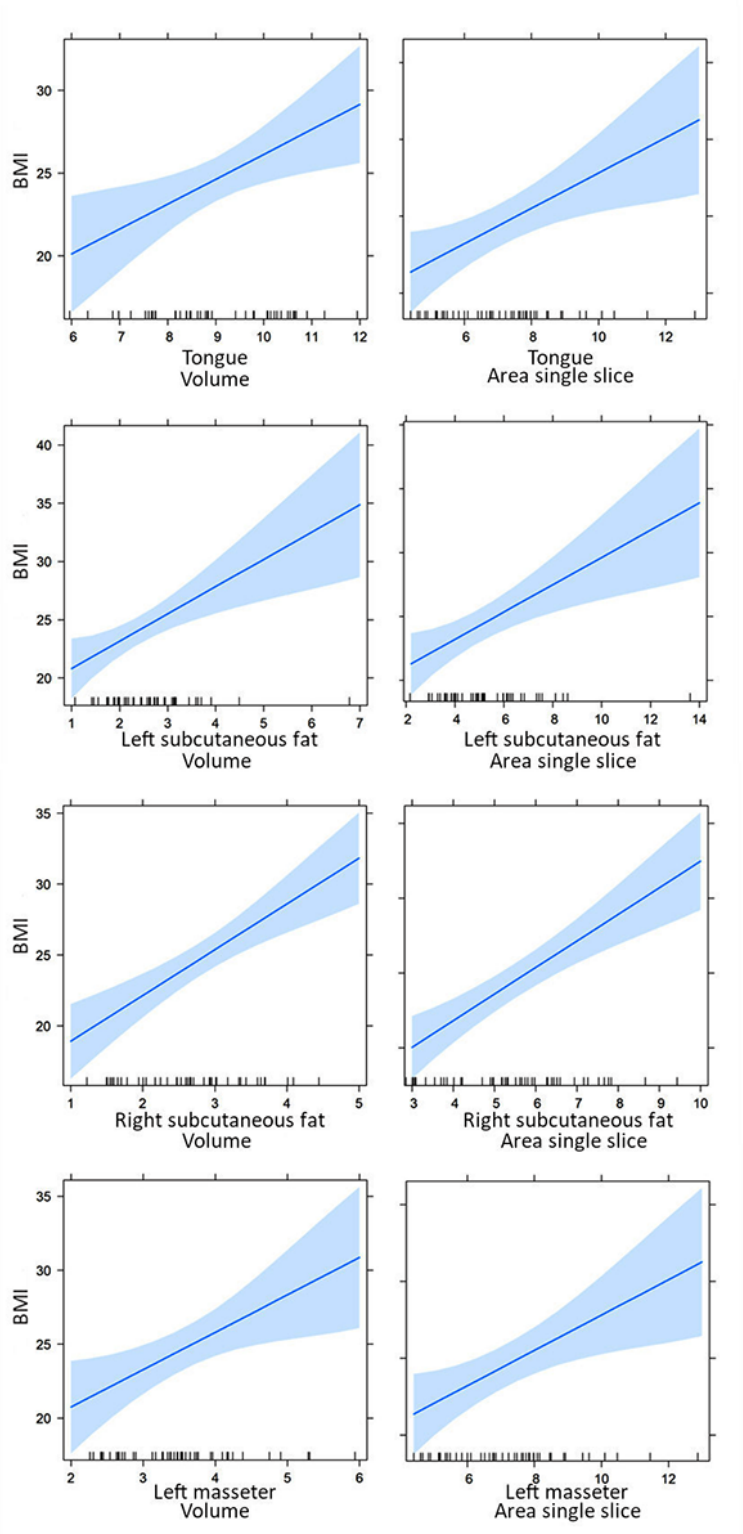
[f]DSC: Dice similarity coefficient.

[g]ASSD: average symmetric surface distance.

Clinical Validation

To prove the clinical validity of the measurements, we evaluated the association between the segmented muscles and subcutaneous fat and BMI. We found a significant positive association between tongue muscle, left masseter muscle, and left and right subcutaneous fat and BMI (Figure 3). The area of a single slice as well as the volume of 5 slices per patient was calculated for this experiment. The results were adjusted by age and sex with $P<.05$ (Multimedia Appendix 2).

Figure 3. Individual linear regression repressing the relationship between quantitative results (area of a single slice and volume of 5 slices) from the Dense U-Net model and BMI.



Discussion

In this study, we evaluated the performance of six deep learning models for segmentation of the masseter muscles, subcutaneous fat, and tongue muscle in MRI images of the head. Several variations of the U-Net architecture were trained and tested using k-fold cross-validation. The use of deep neural networks for segmenting musculoskeletal tissues in patients with AD and LBD is a novel experimental contribution to the deep learning-based segmentation literature as well as the clinical literature.

Our study demonstrated that the Dense U-Net model performed superiorly to the other models in all evaluated regions by achieving an overall DSC of 93.43% and ASSD of 0.68 mm on the test data. The remaining models demonstrated comparable outcomes, except the Attention U-Net, which achieved slightly

less accurate results in all regions with an overall DSC of 92.46% and ASSD of 0.79 mm on the test data. Notably, despite having a comparable number of trainable parameters, the Dense U-Net model demonstrated a higher DSC and lower ASSD than the Wide U-Net model. The Dense U-Net model achieves superior performance with a similar number of parameters, suggesting that its architectural efficiency, rather than parameter quantity alone, drives this improvement.

The validation set produced similar results to the training set, indicating that the trained models did not suffer from overfitting. Although we applied data augmentation to the training data set (results not shown), it did not significantly improve the accuracy of the segmentation models.

Furthermore, we observed a significant correlation between the results and BMI, a well-established measure of nutrition and body composition. This underscores the validity and clinical relevance of this method. If these localized muscle measurements correlate with BMI, it suggests that they may reflect broader nutritional and body composition states.

Sarcopenia is a relatively newly recognized condition for which neuromuscular degeneration, central nervous system alpha motor unit loss, and fat infiltration into muscle are the most distinctive proven and observed pathogenic features, leading to loss of muscle mass and strength and predisposing to physical frailty [2]. We have proposed that non-invasive assessment of intermuscular adipose tissue and muscle mass by image analysis could constitute a viable method to diagnose sarcopenia and predict its associated outcomes, the clinical impact of which is also under study by our group [29].

The validity and clinical implications of measuring the masseter muscle have been shown in previous studies [30,31]. In recent work by our group, we compared the diagnostic capacity for sarcopenia between the gold-standard dual X-ray absorptiometry and our measurements of head muscles. The results showed that both methods had equivalent accuracy [32]. In older adults with glioblastoma, a decreased masseter diameter on preoperative imaging was associated with shorter overall survival and 90-day mortality after surgical resection [33]. In addition, low masseter muscle was significantly associated with worse overall survival in patients aged 65 years or older, diagnosed with squamous cell carcinoma of the head and neck and treated with curative intent [34]. Another study evaluated post-operative results after carotid endarterectomy; low masseter mass was associated with a prolonged hospital stay and recurrent stroke within 5 years [35]. In another study, preoperative masseter mass was a predictor of postoperative pneumonia in patients with esophageal cancer [36]. Additionally, other studies have shown that the masseter muscle can be used as a nutritional biomarker. The masseter muscle, analyzed via computed tomography (CT) anthropometry, showed a statistically significant association with systemic nutritional biomarkers [37]. On the other hand, the tongue has shown to be a good marker of prognosis, as tongue strength has shown to be helpful in diagnosing sarcopenia [38]. Previous studies from our group also report that tongue muscle volume is correlated with malnutrition and even brain structures in patients with dementia [18,29].

Therefore, the approach we present in this paper can be opportunistically used to quantify muscle volumes and investigate the implications of having low muscle mass in the masseter or the tongue in people with brain, head, and neck diseases. Thus, it is an important first step toward developing a more efficient method to estimate masseter and tongue muscle with a better capacity to be implemented in clinical practice. Manual and semi-automatic techniques have been employed in several studies for masseter muscle segmentation in MRI [39]. A recent study used shape determination to segment the masseter muscle in MRI images [40]. In this approach, a manual contour for 8 slices must be defined by the user, and the model then determines the shape for the remaining scans, reducing the time and labor required for segmentation. However, this technique still requires manual segmentation, which can be time-consuming and prone to user error compared to our approach.

Model-based techniques have also been explored for the segmentation of the mastication muscle [41,42]. Although these techniques have shown high accuracy (>90%), there must be a distinct boundary between the masseter muscle and surrounding tissues to ensure accuracy. This distinct boundary refers to visible differences in intensity, texture, or anatomical structure on imaging, which enable the models to accurately separate the muscle from adjacent tissues, such as fat or bone.

Machine and deep learning approaches have been widely used to segment muscles in various body parts. CT scans and cone beam CT scans have primarily been investigated for measuring the head organs, including masseter muscle using these techniques [43-45]. In a previous study, a basic U-Net model was applied to segment the masseter muscle in head CT scans to investigate hemifacial microsomia [46]. The mean DSC reported in that study was 79.4% for the experimental group and 82.4% for the group with mandible deformities, which are lower compared to the results obtained in our study. In a study using CT scan of the head for segmentation of masticatory muscles, deep learning techniques were superior to atlas-based techniques, achieving a mean DSC of 83% [47]. To the best of our knowledge, the techniques considered in our study have not yet been used to segment musculoskeletal tissues in MRI images of the head.

Our study had some limitations. The Attention U-Net model demonstrated the lowest DSC values and the highest ASSD values among the evaluated models, indicating suboptimal segmentation accuracy across tissues. This underperformance may be attributed to the Attention mechanism's inability to effectively focus on the target tissue, leading to dispersed attention, particularly in smaller structures or regions with indistinct boundaries. To address this limitation, future work could involve refining the Attention mechanism to enhance its specificity and focus on regions of interest. Alternatively, exploring models that prioritize multi-scale feature extraction and detail preservation may provide improved segmentation performance, particularly for small or complex anatomical regions.

The DemVest study had some limitations that may have impacted the results. Selection bias might have been present if

patients with more complex health conditions were included, as primary care referrals were used. The study was not specifically designed for the paper's purpose, which could limit the data analysis and interaction control. For example, the absence of a healthy control group prevents us from determining whether the observed muscle volume characteristics are specific to individuals with AD or neurocognitive disorders, or if they represent normal variations associated with aging. However, the primary objective of this study was to demonstrate the feasibility and accuracy of our deep learning model in quantifying muscle volumes using head MRI, rather than to establish definitive differences between diseased and healthy populations. In addition, the sample size is relatively small, and there is a risk that our results may not generalize to larger populations. Additionally, the MRI scans were obtained from a single center using a single MRI machine, which may affect the model's generalizability. Longitudinal analyses were not conducted because imaging was only performed at baseline, in line with the initial primary objectives of the DemVest study and the resources allocated for image acquisition.

These limitations should be considered when interpreting this study's results and addressed in future studies seeking broader application of the proposed approach. Whether masseter and tongue volumes in these muscles correlate with lean body mass and inflammaging and could be predictors of neurocognitive conditions and their associated outcomes remains unknown.

On the other hand, this study has several strengths, including well-characterized data, detailed and exhaustive diagnosis to correctly identify people with dementia, and an automated quality control and analysis pipeline. Additionally, the results fill a gap in the literature and provide insights into a possible method to efficiently diagnose sarcopenia in context when a head MRI is already available.

In summary, to our knowledge, this is the first study that validates deep learning methods that could be easily implemented in clinical practice to measure masseter and tongue muscle volumes with a solid potential to become biomarkers with strong predictive value for adverse outcomes in older persons with dementia. Since imaging is widely used in memory clinics worldwide, this opportunistic approach to image analyses could become standard practice in those settings. However, further large longitudinal studies are still required.

Acknowledgments

We want to thank the participants, researchers, and technical staff that made the DemVest study possible, as well as the staff and facilities provided by (SESAM) Centre for age-related medicine at Stavanger, Norway. This work was supported by the Norwegian government, through hospital owner Helse Vest (Western Norway Regional Health Authority) number 911973. It is also funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust, King's College London, and the Australian Institute for Musculoskeletal Science (AIMSS). Mahdi Imani is supported by an Australian Government Research Training Program (TRP) Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table presenting extended accuracy evaluation metrics, including Dice Similarity Coefficient (DSC), Jaccard Coefficient (JC), Average Symmetric Surface Distance (ASSD), Precision, Recall, Sensitivity, Specificity, and F1-Score. Results are presented as mean (standard deviation) and median (interquartile range) across 10-fold cross-validation.
[XLSX File, 20 KB - aging_v8i1e63686_app1.xlsx]

Multimedia Appendix 2
Table presenting association between area and volumes from MRI scans and BMI in each region of study.
[DOCX File, 16 KB - aging_v8i1e63686_app2.docx]

References

1. WHO G. WHO Methods and Data Sources for Global Burden of Disease Estimates 2000– 2011. : Department of Health Statistics and Information Systems; 2020.
2. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 1;48(1):16-31. [doi: 10.1093/ageing/afy169] [Medline: 30312372]
3. Shafiee G, Keshtkar A, Soltani A, Ahadi Z, Larijani B, Heshmat R. Prevalence of sarcopenia in the world: a systematic review and meta- analysis of general population studies. J Diabetes Metab Disord 2017;16:21. [doi: 10.1186/s40200-017-0302-x] [Medline: 28523252]
4. Cruz-Jentoft AJ, Landi F, Schneider SM, et al. Prevalence of and interventions for sarcopenia in ageing adults: a systematic review. Report of the International Sarcopenia Initiative (EWGSOP and IWGS). Age Ageing 2014 Nov;43(6):748-759. [doi: 10.1093/ageing/afu115] [Medline: 25241753]

5. Morley JE. Sarcopenia: diagnosis and treatment. J Nutr Health Aging 2008;12(7):452-456. [doi: 10.1007/BF02982705] [Medline: 18615226]
6. Tan LF, Lim ZY, Choe R, Seetharaman S, Merchant R. Screening for frailty and sarcopenia among older persons in medical outpatient clinics and its associations with healthcare burden. J Am Med Dir Assoc 2017 Jul 1;18(7):583-587. [doi: 10.1016/j.jamda.2017.01.004] [Medline: 28242192]
7. Cobo CS, Pérez V, Hermosilla C, Nuñez M, Lorena PD. Prevalence of sarcopenia in elderly with dementia institutionalized: a multicenter study. J Aging Res 2014;3:178-181. [doi: 10.14283/jarcp.2014.31]
8. Sperlich E, Fleiner T, Zijlstra W, Haussermann P, Morat T. Sarcopenia in geriatric psychiatry: feasibility of the diagnostic process and estimation of prevalence within a hospital context. J Cachexia Sarcopenia Muscle 2021 Oct;12(5):1153-1160. [doi: 10.1002/jcsm.12748] [Medline: 34151538]
9. Peng TC, Chen WL, Wu LW, Chang YW, Kao TW. Sarcopenia and cognitive impairment: a systematic review and meta-analysis. Clin Nutr 2020 Sep;39(9):2695-2701. [doi: 10.1016/j.clnu.2019.12.014] [Medline: 31917049]
10. van Spronsen PH, Weijs WA, Valk J, Prahl-Andersen B, van Ginkel FC. Comparison of jaw-muscle bite-force cross-sections obtained by means of magnetic resonance imaging and high-resolution CT scanning. J Dent Res 1989 Dec;68(12):1765-1770. [doi: 10.1177/00220345890680120901] [Medline: 2600258]
11. Lee YH, Lee KM, Auh QS. MRI-based assessment of masticatory muscle changes in TMD patients after Whiplash injury. J Clin Med 2021 Apr 1;10(7):1404. [doi: 10.3390/jcm10071404] [Medline: 33915742]
12. Tay L, Ding YY, Leung BP, et al. Sex-specific differences in risk factors for sarcopenia amongst community-dwelling older adults. Age (Dordr) 2015 Dec;37(6):121. [doi: 10.1007/s11357-015-9860-3] [Medline: 26607157]
13. Fung FY, Koh YLE, Malhotra R, et al. Prevalence of and factors associated with sarcopenia among multi-ethnic ambulatory older Asians with type 2 diabetes mellitus in a primary care setting. BMC Geriatr 2019 Apr 29;19(1):122. [doi: 10.1186/s12877-019-1137-8] [Medline: 31035928]
14. van Heusden HC, Chargi N, Dankbaar JW, Smid EJ, de Bree R. Masseter muscle parameters can function as an alternative for skeletal muscle mass assessments on cross-sectional imaging at lumbar or cervical vertebral levels. Quant Imaging Med Surg 2022 Jan;12(1):15-27. [doi: 10.21037/qims-21-43] [Medline: 34993057]
15. Aarsland D, Rongve A, Nore SP, et al. Frequency and case identification of dementia with Lewy bodies using the revised consensus criteria. Dement Geriatr Cogn Disord 2008;26(5):445-452. [doi: 10.1159/000165917] [Medline: 18974647]
16. Bell CC. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders. J Am Med Assoc 1994 Sep 14;272(10):828. [doi: 10.1001/jama.1994.03520100096046] [Medline: 7933395]
17. Skogseth R, Hortobágyi T, Soennesyn H, et al. Accuracy of clinical diagnosis of dementia with Lewy bodies versus neuropathology. J Alzheimers Dis 2017;59(4):1139-1152. [doi: 10.3233/JAD-170274] [Medline: 28731443]
18. Borda MG, Bani Hassan E, Weon JH, et al. Muscle volume and intramuscular fat of the tongue evaluated with MRI predict malnutrition in people living with dementia: a 5-year follow-up study. J Gerontol A Biol Sci Med Sci 2022 Feb 3;77(2):228-234. [doi: 10.1093/gerona/glab224] [Medline: 34338751]
19. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods 2019 Jan;16(1):67-70. [doi: 10.1038/s41592-018-0261-2] [Medline: 30559429]
20. Ronneberger O, Fischer P, Brox T, editors. U-net: convolutional networks for biomedical image segmentation. Presented at: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference; Oct 5-9, 2015; Munich, Germany.
21. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: learning where to look for the pancreas. arXiv. Preprint posted online on 2018arXiv:180403999.
22. Cai S, Tian Y, Lui H, Zeng H, Wu Y, Chen G. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. Quant Imaging Med Surg 2020 Jun;10(6):1275-1285. [doi: 10.21037/qims-19-1090] [Medline: 32550136]
23. Zhang ZX, Liu QJ, Wang YH. Road extraction by deep residual U-Net. IEEE Geosci Remote Sensing Lett 2018 May;15(5):749-753. [doi: 10.1109/LGRS.2018.2802944]
24. Punn NS, Agarwal S. Inception U-Net architecture for semantic segmentation to identify nuclei in microscopy cell images. ACM Trans Multimedia Comput Commun Appl 2020 Feb 29;16(1):1-15. [doi: 10.1145/3376922]
25. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging 2020 Jun;39(6):1856-1867. [doi: 10.1109/TMI.2019.2959609] [Medline: 31841402]
26. Zhang ZL, Sabuncu MR. Generalized cross entropy loss for training deep neural networks with noisy labels. In: Adv Neural Inf Process Syst 2018, Vol. 31. [Medline: 000461852003034]
27. Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Garcia-Rodriguez J. A survey on deep learning techniques for image and video semantic segmentation. Appl Soft Comput 2018 Sep;70:41-65. [doi: 10.1016/j.asoc.2018.05.018]
28. Aldoj N, Biavati F, Michallek F, Stober S, Dewey M. Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net. Sci Rep 2020 Aug 31;10(1):14315. [doi: 10.1038/s41598-020-71080-0] [Medline: 32868836]

29. Borda MG, Castellanos-Perilla N, Tovar-Rios DA, Ferreira D, Duque G, Aarsland D. Tongue muscle mass is associated with total grey matter and hippocampal volumes in Dementia with Lewy Bodies. Arch Gerontol Geriatr 2022;100:104647. [doi: 10.1016/j.archger.2022.104647] [Medline: 35134612]
30. Uhlich R, Hu P. Sarcopenia diagnosed using masseter muscle area predictive of early mortality following severe traumatic brain injury. Neural Regen Res 2018 Dec;13(12):2089-2090. [doi: 10.4103/1673-5374.241451] [Medline: 30323130]
31. Wallace JD, Calvo RY, Lewis PR, et al. Sarcopenia as a predictor of mortality in elderly blunt trauma patients: comparing the masseter to the psoas using computed tomography. J Trauma Acute Care Surg 2017 Jan;82(1):65-72. [doi: 10.1097/TA.0000000000001297] [Medline: 27820555]
32. Borda MG, Duque G, Pérez-Zepeda MU, et al. Using magnetic resonance imaging to measure head muscles: an innovative method to opportunistically determine muscle mass and detect sarcopenia. J Cachexia Sarcopenia Muscle 2024 Feb;15(1):189-197. [doi: 10.1002/jcsm.13362] [Medline: 38050325]
33. Morshed RA, Young JS, Casey M, et al. Sarcopenia diagnosed using masseter muscle diameter as a survival correlate in elderly patients with glioblastoma. World Neurosurg 2022 May;161:e448-e463. [doi: 10.1016/j.wneu.2022.02.038] [Medline: 35181534]
34. McGoldrick D, Alsabbagh AY, Pettit L, Shaikh M, Bhatia S. Masseter muscle defined sarcopenia and survival in head and neck cancer patients at a single cancer centre. Br J Oral Maxillofac Surg 2020 Dec;58(10):e157-e158. [doi: 10.1016/j.bjoms.2020.10.073]
35. Hogenbirk RNM, Banning LBD, Visser A, et al. Association between masseter muscle area and thickness and outcome after carotid endarterectomy: a retrospective cohort study. J Clin Med 2022 May 30;11(11):3087. [doi: 10.3390/jcm11113087] [Medline: 35683474]
36. Kamada T, Ohdaira H, Ito E, et al. Association between masseter muscle sarcopenia and postoperative pneumonia in patients with esophageal cancer. Sci Rep 2022 Sep 30;12(1):16374. [doi: 10.1038/s41598-022-20967-1] [Medline: 36180776]
37. Hwang Y, Lee YH, Cho DH, Kim M, Lee DS, Cho HJ. Applicability of the masseter muscle as a nutritional biomarker. Medicine (Baltimore) 2020 Feb;99(6):e19069. [doi: 10.1097/MD.0000000000019069] [Medline: 32028430]
38. Chen KC, Lee TM, Wu WT, Wang TG, Han DS, Chang KV. Assessment of tongue strength in sarcopenia and sarcopenic dysphagia: a systematic review and meta-analysis. Front Nutr 2021;8:684840. [doi: 10.3389/fnut.2021.684840] [Medline: 34249993]
39. Volk GF, Karamyan I, Klingner CM, Reichenbach JR, Guntinas-Lichius O. Quantitative magnetic resonance imaging volumetry of facial muscles in healthy patients with facial palsy. Plast Reconstr Surg Glob Open 2014 Jun;2(6):e173. [doi: 10.1097/GOX.0000000000000128] [Medline: 25289366]
40. Ng HP, Foong KWC, Ong SH, et al. Quantitative analysis of human masticatory muscles using magnetic resonance imaging. Dentomaxillofac Radiol 2009 May;38(4):224-231. [doi: 10.1259/dmfr/75198413] [Medline: 19372110]
41. Ng HP, Ong SH, Hu Q, Foong KWC, Goh PS, Nowinski WL. Muscles of mastication model-based MR image segmentation. Int J CARS 2006 Nov 23;1(3):137-148. [doi: 10.1007/s11548-006-0046-4]
42. Ng HP, Ong SH, Liu J, et al. 3D segmentation and quantification of a masticatory muscle from MR data using patient-specific models and matching distributions. J Digit Imaging 2009 Oct;22(5):449-462. [doi: 10.1007/s10278-008-9132-1] [Medline: 18516642]
43. Jiang Y, Shang F, Peng J, et al. Automatic masseter muscle accurate segmentation from CBCT using deep learning-based model. J Clin Med 2022 Dec 21;12(1):55. [doi: 10.3390/jcm12010055] [Medline: 36614860]
44. Zhang Y, Pei Y, Qin H, et al. Masseter muscle segmentation from cone-beam ct images using generative adversarial network. In: Zha H, editor. Presented at: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI); Apr 8-11, 2019; Venice, Italy. [doi: 10.1109/ISBI.2019.8759426]
45. Zhu W, Huang Y, Zeng L, et al. AnatomyNet: deep learning for fast and fully automated whole-volume segmentation of head and neck anatomy. Med Phys 2019 Feb;46(2):576-589. [doi: 10.1002/mp.13300] [Medline: 30480818]
46. Han W, Xia W, Zhang Z, et al. Radiomics and artificial intelligence study of masseter muscle segmentation in patients with hemifacial microsomia. J Craniofac Surg 2023;34(2):809-812. [doi: 10.1097/SCS.0000000000009105] [Medline: 36728424]
47. Chen W, Li Y, Dyer BA, et al. Deep learning vs. atlas-based models for fast auto-segmentation of the masticatory muscles on head and neck CT images. Radiat Oncol 2020 Jul 20;15(1):176. [doi: 10.1186/s13014-020-01617-0] [Medline: 32690103]

Abbreviations

AD: Alzheimer disease
ASSD: average symmetric surface distance
CSE: categorical cross-entropy
CT: computed tomography
DLB: dementia with Lewy bodies
DSC: Dice similarity coefficient
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
GT: ground truth

MRI: magnetic resonance imaging

Edited by M Gray; submitted 26.06.24; peer-reviewed by Q Gang, Z Zhao; revised version received 17.12.24; accepted 02.01.25; published 19.03.25.

Please cite as:
Imani M, Borda MG, Vogrin S, Meijering E, Aarsland D, Duque G
Using Deep Learning to Perform Automatic Quantitative Measurement of Masseter and Tongue Muscles in Persons With Dementia: Cross-Sectional Study
JMIR Aging 2025;8:e63686
URL: https://aging.jmir.org/2025/1/e63686
doi:10.2196/63686

Exploring Older Adults' Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study

Arkers Kwan Ching Wong[1], PhD; Jessica Hiu Toon Lee[1], BSc; Yue Zhao[2], PhD; Qi Lu[2], PhD; Shulan Yang[3], MSc; Vivian Chi Ching Hui[1], PhD

[1]School of Nursing, The Hong Kong Polytechnic University, 1 Yuk Choi Rd, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

[2]School of Nursing, Tianjin Medical University, Tianjin, China

[3]Nursing Department, Zhejiang Hospital, Hangzhou, China

Corresponding Author:
Arkers Kwan Ching Wong, PhD
School of Nursing, The Hong Kong Polytechnic University, 1 Yuk Choi Rd, Hung Hom, Kowloon, Hong Kong, China (Hong Kong)

Abstract

Background: Artificial intelligence (AI) is increasingly being applied in various health care services due to its enhanced efficiency and accuracy. As the population ages, AI-based health technologies could be a potent tool in older adults' health care to address growing, complex, and challenging health needs. This study aimed to investigate perspectives on and acceptability of the use of AI-led health technologies among older adults and the potential challenges that they face in adopting them. The findings from this inquiry could inform the designing of more acceptable and user-friendly AI-based health technologies.

Objective: The objectives of the study were (1) to investigate the attitudes and perceptions of older adults toward the use of AI-based health technologies; (2) to identify potential facilitators, barriers, and challenges influencing older adults' preferences toward AI-based health technologies; and (3) to inform strategies that can promote and facilitate the use of AI-based health technologies among older adults.

Methods: This study adopted a qualitative descriptive design. A total of 27 community-dwelling older adults were recruited from a local community center. Three sessions of semistructured interviews were conducted, each lasting 1 hour. The sessions covered five key areas: (1) general impressions of AI-based health technologies; (2) previous experiences with AI-based health technologies; (3) perceptions and attitudes toward AI-based health technologies; (4) anticipated difficulties in using AI-based health technologies and underlying reasons; and (5) willingness, preferences, and motivations for accepting AI-based health technologies. Thematic analysis was applied for data analysis. The Theoretical Domains Framework and the Capability, Opportunity, Motivation, and Behavior (COM-B) model behavior change wheel were integrated into the analysis. Identified theoretical domains were mapped directly to the COM-B model to determine corresponding strategies for enhancing the acceptability of AI-based health technologies among older adults.

Results: The analysis identified 9 of the 14 Theoretical Domains Framework domains—knowledge, skills, social influences, environmental context and resources, beliefs about capabilities, beliefs about consequences, intentions, goals, and emotion. These domains were mapped to 6 components of the COM-B model. While most participants acknowledged the potential benefits of AI-based health technologies, they emphasized the irreplaceable role of human expertise and interaction. Participants expressed concerns about the usability of AI technologies, highlighting the need for user-friendly and tailored AI solutions. Privacy concerns and the importance of robust security measures were also emphasized as critical factors affecting their willingness to adopt AI-based health technologies.

Conclusions: Integrating AI as a supportive tool alongside health care providers, rather than regarding it as a replacement, was highlighted as a key strategy for promoting acceptance. Government support and clear guidelines are needed to promote ethical AI implementation in health care. These measures can improve health outcomes in the older adult population by encouraging the adoption of AI-driven health technologies.

(JMIR Aging 2025;8:e66778) doi:10.2196/66778

KEYWORDS

artificial intelligence–based health technologies; health technology; AI-based health technology; machine learning; ML; artificial intelligence; AI; algorithm; model; analytics; perceptions; acceptability; gerontology; geriatrics; older adult; elderly; older person; older people; aging; mobile phone

Introduction

Background

Artificial intelligence (AI) refers to a computerized system that is capable of executing a wide range of tasks that typically involve human intelligence—ranging from physical tasks, cognitive functions, and problem-solving to decision-making—and can be performed without explicit instructions from humans [1]. AI can be classified into 4 main subsets—machine learning, natural language processing (NLP), physical robots, and robotic process automation. The field of AI is rapidly advancing, and AI technologies have been widely applied in different aspects of health care services, including diagnostic assistance, health screening, and imaging interpretations [2]. It has been described as a "second set of eyes" for medical practitioners [3]. Furthermore, over the past decade, the application of AI systems has increasingly centered on empowering patients to actively manage their health and boost their participation in the shared decision-making process, giving rise to diverse AI-driven products such as robots, smart assistants, virtual or augmented reality, wearable devices, and mobile apps [4]. Particularly notable is the emergence of AI-powered chatbot services, which permit patients to seek medical advice and receive triage for their conditions in a prompt and cost-effective manner [5]. With the debut of ChatGPT by OpenAI in November 2022 and the popularity it has gained, it is envisioned that AI-based conversational large language models with NLP abilities could conceivably revolutionize health care practice and education, contributing to substantial transformative shifts [6,7].

The world is currently encountering a significant expansion in the aging population, with the number of people aged 60 years or above anticipated to rise to 1.4 billion by 2030 and 2.1 billion by 2050 [8]. Conventional older adult health care has relied heavily on in-person monitoring; however, the further intensification of the shortage of health care workers could in the long run present a global challenge to the sustainability of delivering quality medical care [9]. In recent years, digital technologies, such as smart older adult health care products, which involve the deep integration of AI, have been used in older adult health care sectors [10]. Evidence has shown that AI-based technologies can play a constructive role in improving the physical and psychological well-being, quality of life, and independent living of older adults [11]. However, there is a great digital divide between the older and younger generations, with the former being the age group that is the least likely to have access to computers and the internet, due to physical obstacles such as physical disabilities and to psychological factors such as a lack confidence in using technology as well as ethical concerns [12,13].

While AI-based health technologies have rapidly evolved, and some studies have comprehensively discussed the advantages of applying them in the field of geriatrics, insufficient academic attention has been paid to acquiring a clear understanding of the attitudes, perceptions, and experiences of seniors toward these technologies [10]. This gap may implicitly influence the acceptance of, and motivation to use, these technologies in the future. In addition, while studies have primarily focused on the perceptions of medical practitioners and patients regarding AI health technologies, the perspectives of older adults—a key user group—have received limited attention [14-16]. For instance, although some research has explored older adults' views on general AI-powered technologies, their relevance to AI health technologies remains uncertain [17]. Concerns about accuracy, reliability, and trust in AI tools compared with in-person medical advice further underscore the need for targeted investigation. This study addresses this gap by exploring older adults' perceptions of and acceptability toward AI-based health technologies, with a specific focus on practical strategies to enhance adoption. By identifying barriers, facilitators, and preferences, the findings have clear clinical implications, offering actionable insights for health care providers, policy makers, and AI developers. These insights can inform the design of tailored, user-friendly AI tools and guide their ethical implementation in health care, ultimately improving older adults' health outcomes.

Theoretical Framework

The interview guide for this study was developed using the Theoretical Domains Framework (TDF) and the Capability, Opportunity, Motivation, and Behavior (COM-B) model behavior change wheel to comprehensively explore factors influencing the acceptability of AI-based health technologies among older adults. The TDF consists of 14 domains that integrate behavioral determinants derived from over 30 psychological theories [18]. To ensure comprehensive coverage of these domains, the interview guide focused on five key areas: (1) general impressions of AI, (2) previous experiences with AI-based technologies, (3) perceptions and attitudes toward AI-based health technologies, (4) expected difficulties and underlying reasons, and (5) willingness. These focus areas were aligned with the study objectives and systematically mapped to the relevant TDF domains, as summarized in Table 1. This approach ensured that all 14 domains were addressed, either directly or indirectly, during the interviews.

Table . The interview questions guided by the Theoretical Domains Framework (TDF).

Focus area	Aligned TDF domains	Example question
General impressions of AI[a]	Knowledge, emotion, and optimism	"What comes to mind when you think about artificial intelligence and its role in health technologies?"
Previous experiences with AI-based technologies	Memory, attention, and decision processes; knowledge; and skills	"Have you used any AI-based technologies before? Can you describe your experience?"
Perceptions and attitudes toward AI-based health technologies	Social or professional role and identity; beliefs about capabilities; beliefs about consequences; emotion; and optimism	"What are your thoughts about using AI-based technologies for managing your health?"
Expected difficulties and underlying reasons	Environmental context and resources; skills; social influences; behavioral regulation; and memory, attention, and decision processes	"What challenges do you think you might face when using AI-based health technologies? Why do you think these occur?"
Willingness	Intentions; goals; beliefs about capabilities; and reinforcement	"Would you be willing to use AI-based health technologies? Why or why not?"

[a]AI: artificial intelligence.

To further enhance the applicability of findings, the identified TDF domains were mapped to the COM-B behavior change wheel (Table 2). The COM-B model summarizes 6 sources of behavior—social or physical opportunity, automatic or reflective motivation, and physical or psychological capability. This model provided a practical framework for identifying specific behavior change factors and strategies to improve the acceptance and adoption of AI-based health technologies among older adults [19].

Table . Capability, Opportunity, Motivation, and Behavior (COM-B) model components and its relation to Theoretical Domains Framework (TDF) domains.

COM-B components		TDF domains
Capability		
	Psychological capability	• Knowledge • Skills • Memory, attention, and decision processes
	Physical capability	• Behavioral regulation • Skills
Opportunity		
	Social opportunity	• Social influences
	Physical opportunity	• Environmental context and resources
Motivation		
	Reflective motivation	• Social or professional role and identity • Beliefs about capabilities • Optimism • Beliefs about consequences • Intentions • Goals
	Automatic motivation	• Social or professional role and identity • Optimism • Reinforcement • Emotion

The interview guide included open-ended questions designed to elicit rich, detailed responses.

This structured mapping ensured that the interview guide covered all 14 TDF domains comprehensively, while tailoring the questions to elicit insights into factors influencing the acceptability of AI-based health technologies among older adults. While alternative models, such as the technology acceptance model and the unified theory of acceptance and use of technology are commonly used in technology adoption studies, these models primarily focus on individual perceptions, such as perceived usefulness and ease of use (technology acceptance model) or performance expectancy and effort expectancy (unified theory of acceptance and use of technology).

These constructs, while valuable, may not fully capture the complex interplay of factors influencing older adults' adoption of AI technologies.

The systematic alignment between the interview guide and the TDF framework facilitated the identification of theoretical mediators and behavioral determinants. This mapping subsequently informed the integration of the findings with the COM-B behavior change wheel, allowing for the development of tailored strategies to promote the use of AI-based health technologies in this population.

Objectives

The objectives of the study were (1) to investigate the attitudes and perceptions of older adults toward the use of AI-based health technologies; (2) to identify potential facilitators, barriers, and challenges influencing older adults' preferences toward AI-based health technologies; and (3) to inform strategies that can promote and facilitate the use of AI-based health technologies among older adults.

Methods

Study Design

A qualitative descriptive design was adopted for this study [20]. This approach helps in the effort to uncover and understand the experiences, attitudes, and perceptions of people, making it appropriate for this study. The findings from a qualitative descriptive study are able to inform strategies that promote and facilitate the use of AI-based health technologies, which makes it particularly useful for this research. To explore perceptions on and acceptability of the use of AI-based technologies in health maintenance among older adults, and to provide insights into their subjective views, in-depth semistructured interviews guided by an interview guide were conducted.

Participants and Recruitment

Community-dwelling older adults aged 60 years or above were recruited from a local community center using a convenience sampling method. Eligible participants included those with or without previous exposure to AI tools who were willing to participate and share their experiences. Exclusion criteria included underlying physical, psychological, or neurodevelopmental problems that impaired interaction, mental instability, cognitive impairments such as traumatic brain injury, substance abuse, dementia, severe psychiatric disorders, intellectual disability, or being bedridden.

Sampling and Sample Size

Convenience sampling approach was employed, guided by the principle of data saturation. The target sample size was 25 participants, which is commonly regarded as sufficient for qualitative descriptive studies to achieve thematic depth and richness [21]. Ultimately, 27 older adults were recruited, providing adequate data to explore the study's focus areas.

Data Collection

Data collection involved 3 sessions of semistructured interviews with each participant, each lasting 1 hour. To minimize ambiguity, AI was classified into 4 categories: machine learning, NLP, physical robots, and robotic process automation. Interviews covered five focus areas: (1) general impressions of AI; (2) previous experiences with AI-based technologies; (3) perceptions and attitudes toward AI-based health technologies; (4) anticipated difficulties in using AI-based health technologies and their underlying reasons; and (5) willingness, preferences, and motivations in accepting AI-based health technologies.

In the first focus area, examples of AI tools were not initially provided to elicit spontaneous responses. However, recognizing that participants might not be fully aware of everyday interactions with AI, the second focus area included real-life examples to clarify potential misconceptions. Examples of machine learning technologies included fitness tracking devices, such as Google Fitbit and Apple Watch, which monitor physical activity and provide insights into health trends. For NLP, the interviewer referenced virtual assistants like Siri (Apple Inc) and Alexa (Amazon Inc), as well as health care chatbots that answer patient inquiries. Computer vision was illustrated with examples of AI systems that analyze medical images, such as detecting abnormalities in x-rays or CT scans. This approach enabled a nuanced exploration of participants' experiences with general AI technologies versus their attitudes toward AI for health purposes. The data were securely stored in Cloud storage, accessible only to the research team.

Data Analysis

Thematic analysis was used in this study [22]. The interview recordings were first translated into English, transcribed, read, and reread to gain a sense of the participants' experiences. Initial ideas were noted down. Next, the data were coded to extract key insights from the transcript. The codes were reviewed in a weekly discussion with the supervisor to obtain agreement on the interpretations. The underlying subthemes that emerged from the codes were identified and the subthemes were further clustered into themes. The themes, subthemes, and representative quotes were used to present the findings. Data were mapped to TDF to identify underlying theoretical mediators that influenced the perceptions and acceptability to the older adults of the use of AI-based health technologies. Once these theoretical domains were identified, they were mapped to components of the COM-B behavior change wheel that matched theoretical mediators with corresponding key strategies.

Ethical Considerations

The study protocol and procedures were approved by the Ethical Committee of the Hong Kong Polytechnic University (approval HSEARS20230810006) on August 29, 2023, and adhered to the ethical standards of the Declaration of Helsinki. Before starting the study, all participants provided informed consent during a face-to-face interview with the researcher. All the participants were fully informed about the nature and purpose of the study. They were guaranteed the right to withdraw from the study at any time without adverse consequences. The collected data were encrypted and stored in a password-protected database. Although the researchers did not foresee any significant risks associated with the proposed study, the researchers recognized that older adults might experience discomfort or fatigue from sitting through the interviews. To address this, the researchers ensured that participants were given

scheduled breaks throughout the interview. In addition, participants were encouraged to request breaks as needed to ensure that they were comfortable during the entire interview process. Participants who completed the interviews received an HK $100 (US $12.80) supermarket gift voucher to compensate for transportation costs.

Results

Participant Characteristics

The study included 27 participants, with 18 females (66.7%) and 9 males (33.3%). The mean age was 69.44 (SD 6.7) years. Most participants were married (19/27, 70.4%) and had a high school education (14/27, 51.9%). A majority lived with family (19/27, 70.4%) and rated their health as good (14/27, 58.3%). In addition, 66.7% (18/27) reported having chronic diseases. Regarding technology, 51.9% (14/27) felt somewhat comfortable using smartphones or tablets, and 59.3% (16/27) were somewhat familiar with tech products. Over 80% (23/27, 85.2%) of participants had heard of AI health technology, though their familiarity was limited. AI product usage varied, with approximately half of the participants reporting they used AI products often or occasionally, while the other half reported they used them seldom or were unfamiliar with them. Most participants (21/27, 77.8%) believed AI was to some extent helpful in health management, and 66.7% (18/27) had a positive overall impression of AI. Further details about the characteristics of the participants are provided in Table 3.

Table . Participant characteristics.

Characteristic		Total (N=27)
Gender, n (%)		
	Men	9 (33.3)
	Women	18 (66.7)
Age (y)		
	Mean (SD)	69.44 (6.17)
	Median (IQR)	70 (66-74)
Marital status, n (%)		
	Single	1 (3.7)
	Married	19 (70.4)
	Divorced	1 (3.7)
	Widowed	6 (22.2)
Education level, n (%)		
	Primary school and below	4 (14.8)
	Junior high school	7 (25.9)
	High school	14 (51.9)
	College or university	2 (7.4)
	Postgraduate degree	0 (0)
Living status, n (%)		
	Live with family	19 (70.4)
	Live alone	7 (25.9)
	Did not complete	1 (3.7)
Self-rated health status, n (%)		
	Poor	1 (3.7)
	Fair	12 (44.4)
	Good	14 (58.3)
Any chronic diseases, n (%)		
	Yes	18 (66.7)
	No	9 (33.3)
Comfort level with using smartphone or tablet, n (%)		
	Very comfortable	10 (37.0)
	Somewhat comfortable	14 (51.9)
	Not very comfortable	2 (7.4)
	Completely uncomfortable	0 (0)
	Did not complete	1 (3.7)
Familiarity with technology products, n (%)		
	Very familiar	0 (0)
	Somewhat familiar	16 (59.3)
	Not very familiar	8 (29.6)
	Completely unfamiliar	2 (7.4)
	Did not complete	1 (3.7)
Have heard of AI [a] health technology, n (%)		
	Yes, very familiar	0 (0)

Characteristic		Total (N=27)
	Yes, somewhat familiar	11 (40.7)
	Yes, not very familiar	12 (44.4)
	No, completely unfamiliar	3 (11.1)
	Did not complete	1 (3.7)
Frequency of AI product usage, n (%)		
	Often used	6 (22.2)
	Occasionally used	6 (22.2)
	Seldom used	5 (18.5)
	Never used	7 (25.9)
	Did not complete	3 (11.1)
Attitude toward AI's helpfulness in health management, n (%)		
	Very helpful	2 (7.4)
	To some extent helpful	21 (77.8)
	Not helpful	0 (0)
	Uncertain	2 (7.4)
	Did not complete	2 (7.4)
Overall impression about AI technology, n (%)		
	Positive	18 (66.7)
	Neutral	6 (22.2)
	Negative	0 (0)
	Uncertain	1 (3.7)
	Did not complete	2 (7.4)

[a]AI: artificial intelligence.

Objective 1: Investigate Older Adults' Attitudes and Perceptions Related to the Use of AI-Based Health Technologies

General Impressions of AI-Based Health Technologies

Most participants had positive views of AI-based health technologies, recognizing their potential to improve health outcomes by enhancing health monitoring, providing personalized recommendations, and assisting in decision-making. However, there was also limited awareness and understanding of specific AI-based health technologies, with some participants confusing them with general technologies such as smartphones. In addition, concerns were raised about fraud and scams associated with AI-based health technologies.

Attitudes and Perceptions Toward AI-Based Health Technologies

Privacy and data security were recurring themes in participants' discussions of AI-based health technologies. Many participants expressed concerns about the potential misuse of personal data and highlighted the importance of robust security measures. These concerns were particularly pronounced among participants with limited experience using technology or those influenced by media reports of data breaches. However, a subset of participants viewed privacy as a societal issue rather than a specific risk associated with AI-based technologies. For instance, one participant remarked:

> *Yes, it's acceptable.…Uh, what privacy do you have? There is no privacy in the whole world. Everyone's mobile phone is being monitored.* [Participant 3]

Another noted:

> *Not to mention privacy nowadays, from where we sit now there is completely no privacy…even if you listen to this song now, then this song will automatically appear (on your phone). When you are viewing clothes, then you will see the clothes later on (on your phone).* [Participant 5]

These contrasting perspectives highlight the varying levels of concern about privacy among older adults and underscore the need for transparent communication about data security in the development and deployment of AI-based health technologies.

Trustworthiness and Accuracy

Trust in AI-based health technologies varied among participants, influenced by previous experiences, information sources, and perceptions of accuracy. While some acknowledged the utility of AI in self-monitoring, others expressed reservations about its accuracy and trustworthiness, stressing the irreplaceable value of human judgment and expertise in health care.

Objective 2: Identify Potential Facilitators, Barriers, and Challenges That Influence Older Adults' Preferences Toward AI-Based Health Technologies

Technological Barriers

Participants identified multiple technological challenges that hindered their ability to use AI-based health technologies effectively. These challenges included operational difficulties, limited digital literacy, and lack of access to required resources, such as hardware or stable internet connectivity. These barriers were particularly evident among participants with physical or cognitive impairments due to age-related degeneration.

For instance, 1 participant described how physical limitations affected their interaction with technology:

> *I have degeneration, too. I could remember a lot of things before. But now I can't. My body movements are slow now, so it's difficult for me to use AI-based health technologies.* [Participant 3]

Another significant barrier was self-perceived incompetence in using digital tools, often leading to frustration or feelings of inadequacy:

> *"I feel like I am stupid after learning something...."* [Participant 7]

In addition, resource-related challenges, such as a lack of access to necessary hardware (eg, smartphones) and stable internet connections, were raised, further limiting participants' ability to engage with AI technologies.

Emotional and Psychological Barriers

Emotional and psychological factors emerged as significant barriers to the adoption of AI-based health technologies. These included fear of technology, skepticism regarding its reliability, and concerns about losing human connection in health care interactions. The prevalence of scams and fraudulent activities in the digital age further compounded participants' fears, making them hesitant to trust digital tools.

For example, 1 participant expressed a deep apprehension about interacting with technology:

> *I don't feel comfortable with these technologies because they seem complicated, and I worry about making mistakes.* [Participant 6]

Concerns about scams were also prevalent, with another participant noting:

> *Nowadays, you hear so much about fraud. It makes me scared to trust anything online, even if it looks helpful.* [Participant 4]

These findings illustrate that emotional and psychological barriers are not only about individual fears but also reflect broader societal concerns regarding trust and safety in the digital world.

Perceived Usefulness and Relevance

Participants perceived AI-based health technologies as useful, particularly those that addressed specific health needs such as the monitoring of blood pressure, blood sugar, and cholesterol. They also appreciated technologies that could assist with health monitoring, reminders, and personalized care.

> *I hope that these three things could be more mature, help to reduce blood pressure, blood sugar, or blood cholesterol, whatever. At least I know whether my blood sugar is good or not; if it is not good, I will eat something and exercise to increase my physical capacity.* [Participant 3]

> *Yeah, like an alarm. As a reminder to alert you. Well, high blood pressure, high blood sugar, high cholesterol level, those we get when are older. If AI can help with these three things, it is definitely good.* [Participant 2]

Acknowledgment by Authorities

The endorsement of AI-based health technologies by official authorities or regulatory bodies significantly facilitated their acceptance and adoption. Participants felt more confident and trusting when these technologies were validated by reputable sources, such as health care agencies or government bodies.

> *If it is something organized by the Hospital Authority or the government, it will be more trustworthy. Because sometimes, if it is made by pharmaceutical companies, it may not necessarily be that clear.* [Participant 2]

> *Well, if you talk about it being official, like the doctors and government, I will assume that this AI is reliable, you will basically feel at ease.* [Participant 1]

> *I have to see if the AI comes from a large or a small company. For example, Hong Kong Polytechnic University, Baptist University — they are more reliable then.* [Participant 4]

Objective 3: Inform Strategies That Promote and Facilitate the Use of AI-Based Health Technologies Among Older Adults

Integration of Human Expertise

While recognizing the potential benefits of AI, participants emphasized the need for AI technologies to complement rather than replace human expertise in health care. They suggested that AI could be useful as an auxiliary tool for making preliminary medical suggestions, but that the final decision-making should involve consultations with a human being.

> *I just ask the AI, but I can ignore its answers or I will think about it again. It is just a reference and I won't believe it completely. Well, when you see a doctor, you won't just see one doctor, you will see many many right?...AI is just the first contact point.* [Participant 2]

User-Friendly Design

To enhance the acceptability of AI-based health technologies, it is crucial to develop solutions that are user-friendly and tailored to the specific needs of older adults. A user-friendly design should include simplified and intuitive interfaces that

minimize cognitive load, with features such as large, high-contrast text and icons for better visibility and ease of use. Voice command functionality can further improve accessibility for users with limited dexterity or vision impairments. In addition, providing step-by-step tutorials, user manuals in multiple formats (eg, videos and print), and accessible customer support can address challenges related to digital literacy. These features collectively ensure that the technology is not only easy to use but also aligns with the physical and cognitive capabilities of older adults, thereby fostering greater adoption and satisfaction.

Addressing Privacy and Security Concerns

Given the concerns around privacy and data security, the development of AI-based health technologies should include robust security measures to protect the personal information of users. Addressing these concerns is essential to building trust and encouraging adoption among older adults.

Government and Regulatory Support

Participants highlighted the importance of government and regulatory support in promoting the ethical implementation of AI in health care. Clear guidelines and endorsements from authoritative bodies can foster trust and enhance the acceptability of AI-based health technologies among older adults.

Table 4 provides a detailed description of how the identified barriers and facilitators toward the use of AI-based health technologies were mapped to the TDF domains and corresponding COM-B components. These mappings were derived from participants' responses during the interviews. The barriers and facilitators were categorized based on their alignment with the TDF domains and subsequently mapped to the COM-B components to identify actionable strategies for behavior change.

Table . Identified Theoretical Domains Framework (TDF) domains and Capability, Opportunity, Motivation, and Behavior (COM-B) components related to the adoption of artificial intelligence (AI)–based health technologies.

Barrier or facilitator	COM-B components	TDF domains	Illustrative participant quote
Barrier: limited technological skills	Capability	Knowledge and skills	"I find it difficult to use these technologies because my body movements are slow, and I don't know how to start." [P3]
Facilitator: perceived usefulness and relevance	Motivation	Beliefs about consequences, intentions, and goals	"If AI can help monitor my blood pressure or remind me, it would be very helpful." [P2]
Barrier: lack of official recognition or endorsement	Opportunity	Social influences	"If it's recognized by the government or hospitals, I would trust it more." [P1]
Facilitator: confidence in trustworthiness and accuracy	Motivation	Beliefs about capabilities	"If the results are accurate, I can rely on it for some guidance." [P4]
Barrier: privacy and data security concerns	Motivation	Emotion	"I worry about my personal information being stolen if I use this technology." [P5]
Facilitator: AI as an auxiliary tool for initial suggestions	Motivation	Beliefs about consequences and intentions	"I can use AI for some advice, but I still prefer to consult a doctor for the final decision." [P2]

Discussion

Principal Findings

The aim of the study was to provide insights into the perceptions and acceptability of AI-based health technologies among older adults, using the TDF and the COM-B behavior change wheel as theoretical frameworks. These frameworks were instrumental in identifying key behavioral determinants and mapping them to actionable strategies for enhancing AI adoption. In total, 9 of the 14 TDF domains (knowledge, skills, social influences, environmental context and resources, beliefs about capabilities, beliefs about consequences, intentions, goals, and emotion) were identified and mapped to 6 COM-B components (psychological capability, physical capability, social opportunity, physical opportunity, reflective motivation, and automatic motivation). By applying these frameworks, the study provided a structured approach to understanding the factors influencing older adults' behavior toward AI adoption. The responses of the participants shed light on various aspects of the adoption and utilization of AI-based health technologies among older adults. First, the participants expressed a range of attitudes and perceptions toward AI-based health technologies, including curiosity, skepticism, and enthusiasm toward the use of AI in the management and maintenance of health. These perceptions align with the TDF domains of knowledge and beliefs about consequences, which underscore the need to enhance understanding of AI and clearly communicate its benefits to older adults. The interviews provided a deeper understanding of the factors influencing these attitudes and the potential benefits and concerns associated with AI adoption in the context of the participants' health. They shared their perceptions of AI technologies, addressing aspects such as reliability, trustworthiness, and their perception of the impacts on health care outcomes. These insights informed the reflective motivation component of the COM-B framework, highlighting the

importance of fostering trust and reliability in AI-based technologies to enhance adoption.

In addition, the study revealed an interesting finding regarding older adults' preferences for using AI in asking for advice and self-health maintenance. Contrary to expectations, participants expressed a higher level of trust and reliability in real persons, particularly physicians, when making decisions about pharmacological interventions such as changing drug dosages. Participants believed that AI could play a valuable role in providing information and acting as a reference tool for self-health maintenance. However, they expressed reservations about relying solely on AI for decisions related to medication management. This finding is linked to the TDF domain of social influences and the social opportunity component of the COM-B framework, as participants emphasized the irreplaceable role of human expertise and interpersonal trust in health care decision-making. They viewed AI as a useful source of information but considered the expertise and experience of physicians and other health care professionals to be more reliable and crucial in making decisions regarding medical interventions and treatment. The emotion domain of TDF also played a role, as participants' skepticism and emotional barriers highlighted the importance of addressing concerns through education and reassurance. The findings were also in line with a mixed methods study indicating that the lack of empathy and a professional human approach made AI chatbots less acceptable to some users [15]. This insight has significant implications for the development of AI-based health technologies. It might indicate that older adults desire a collaborative approach that combines the benefits of AI with the guidance and expertise of real-person health care providers [23-25]. Developers could focus on creating AI systems that could provide accurate and evidence-based information, acting as a reference tool to support older adults' self-health maintenance efforts. In this context, AI could assist older adults in accessing reliable and up-to-date information about medications, potential side effects, and alternative treatment options. AI-based systems could also offer personalized recommendations for lifestyle modifications, nonpharmacological interventions, and self-care strategies that align with the preferences and needs of older adults.

While privacy and confidentiality are often considered to be potential concerns when using AI in health care, surprisingly, many expressed a relaxed and relatively pessimistic attitude toward privacy. One of the reasons for this was that the perceived benefits outweighed the concerns. Most expressed the view that the potential benefits of AI-based health technologies have overshadowed any concerns about privacy or confidentiality. They may have viewed the potential improvement in their health outcomes or access to personalized care as more significant than the potential risks to their privacy. Another underlying reason may be that the participants may have had other concerns that they considered more important than privacy and confidentiality in the context of AI-based health technologies. For example, they may have been more focused on the usability, effectiveness, or reliability of the technology. Regardless, the absence of expressed concerns does not necessarily indicate the lack of importance of privacy and confidentiality. Rather, it highlights the need for researchers and developers to proactively address these concerns and ensure that robust privacy and security measures are in place when designing and implementing AI-based health technologies. It is crucial for future research and development efforts to emphasize the importance of privacy and confidentiality in AI-based health technologies, which involves implementing strong data protection measures, obtaining informed consent from users, and transparently communicating how their data will be used and safeguarded.

To enhance older adults' acceptance of AI-based health technologies, strategies should focus on improving technological skills through tailored training programs, securing official recognition and endorsement from health care authorities, addressing concerns over privacy and data security with robust protocols, emphasizing the trustworthiness and accuracy of AI tools, positioning AI as auxiliary aids for medical suggestions, highlighting the usefulness and relevance of AI applications, and providing emotional support to mitigate psychological barriers. These strategies are directly informed by the COM-B framework, targeting capability, opportunity, and motivation as the key drivers of behavior change. By implementing these targeted strategies, older adults might be encouraged to embrace AI technologies, leading to improved health care outcomes and increased engagement with innovative health care solutions. Meanwhile, government endorsements could also contribute to the establishment of guidelines, standards, and regulations that ensure the ethical use of AI in health care, addressing concerns related to privacy, data security, and quality assurance.

Limitations

The limitations of this study were the sample size and representativeness of participants. While we recognize that this sample size may not fully capture the diversity of the population, the study aimed to offer initial, in-depth insights into older adults' perceptions of AI-based health technologies. Future studies could address subgroup comparisons by incorporating stratified sampling to examine variations across different demographic or socioeconomic groups. In addition, while quantitative comparisons were not a focus of this study, future research could include mixed method approaches to statistically assess participant perceptions and examine their relationships with demographic and contextual factors by using validated survey tools. Last but not the least, many participants had limited before knowledge or experience with AI-driven health technologies. Some participants conflated AI technologies with general digital tools, such as mobile apps, due to a lack of familiarity with the distinction. To address this, the interviewer provided concise explanations and real-life examples of AI subsets, including machine learning, NLP, computer vision, and expert systems. While this approach helped participants better understand the discussion, their responses were still shaped by their interpretations and familiarity with technology. Consequently, the findings reflect older adults' perceptions and acceptance of health technologies they associate with AI, rather than definitive perspectives on AI-driven health technologies.

Conclusion

While the majority of participants recognized the potential advantages of AI-based health technologies, they underscored

the indispensable role of human expertise and interaction. Vital strategies for improving acceptability include crafting user-friendly and customized AI solutions, addressing privacy concerns, ensuring robust security measures, and integrating AI as a complementary tool alongside health care providers. Government backing and guidelines can play a pivotal role in advancing ethical AI integration in health care, fostering trust, and enhancing personalized care for older adults, ultimately leading to improved health outcomes in this population.

Conflicts of Interest

None declared.

References

1. Esmaeilzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: experimental study. J Med Internet Res 2021 Nov 25;23(11):e25856. [doi: 10.2196/25856] [Medline: 34842535]
2. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Jun 21;2(4):230-243. [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
3. Cai CJ, Winter S, Steiner D, Wilcox L, Terry M. "Hello AI": uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. Proc ACM Hum-Comput Interact 2019 Nov 7;3(CSCW):1-24. [doi: 10.1145/3359206]
4. Zhang Z, Genc Y, Wang D, Ahsen ME, Fan X. Effect of AI explanations on human perceptions of patient-facing AI-powered healthcare systems. J Med Syst 2021 May 4;45(6):64. [doi: 10.1007/s10916-021-01743-6] [Medline: 33948743]
5. Fan X, Chao D, Zhang Z, Wang D, Li X, Tian F. Utilization of self-diagnosis health chatbots in real-world settings: case study. J Med Internet Res 2021 Jan 6;23(1):e19928. [doi: 10.2196/19928] [Medline: 33404508]
6. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare (Basel) 2023 Mar 19;11(6):887. [doi: 10.3390/healthcare11060887] [Medline: 36981544]
7. Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and research? Diabetes & Metabolic Syndrome: Clinical Research & Reviews 2023 Apr;17(4):102744. [doi: 10.1016/j.dsx.2023.102744]
8. Ageing. World Health Organization. URL: https://www.who.int/health-topics/ageing#tab=tab_1 [accessed 2025-01-31]
9. Ho A. Are we ready for artificial intelligence health monitoring in elder care? BMC Geriatr 2020 Sep 21;20(1):358. [doi: 10.1186/s12877-020-01764-9] [Medline: 32957946]
10. Lee CH, Wang C, Fan X, Li F, Chen CH. Artificial intelligence-enabled digital transformation in elderly healthcare field: scoping review. Advanced Engineering Informatics 2023 Jan;55:101874. [doi: 10.1016/j.aei.2023.101874]
11. Czaja SJ, Ceruso M. The promise of artificial intelligence in supporting an aging population. J Cogn Eng Decis Mak 2022 Dec;16(4):182-193. [doi: 10.1177/15553434221129914]
12. Ball C, Francis J, Huang KT, Kadylak T, Cotten SR, Rikard RV. The physical-digital divide: exploring the social gap between digital natives and physical natives. J Appl Gerontol 2019 Aug;38(8):1167-1184. [doi: 10.1177/0733464817732518] [Medline: 29165009]
13. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
14. Fritsch SJ, Blankenheim A, Wahl A, et al. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. Digit Health 2022;8:20552076221116772. [doi: 10.1177/20552076221116772] [Medline: 35983102]
15. Nadarzynski T, Miles O, Cowie A, Ridge D. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. Digit Health 2019;5:2055207619871808. [doi: 10.1177/2055207619871808] [Medline: 31467682]
16. Shinners L, Aggar C, Grace S, Smith S. Exploring healthcare professionals' understanding and experiences of artificial intelligence technology use in the delivery of healthcare: an integrative review. Health Informatics J 2020 Jun;26(2):1225-1236. [doi: 10.1177/1460458219874641] [Medline: 31566454]
17. Shandilya E, Fan M. Understanding older adults' perceptions and challenges in using AI-enabled everyday technologies. Presented at: Chinese CHI 2022: The Tenth International Symposium of Chinese CHI; Oct 22-23, 2022; Guangzhou, China p. 105-116. [doi: 10.1145/3565698.3565774]
18. Richardson M, Khouja CL, Sutcliffe K, Thomas J. Using the theoretical domains framework and the behavioural change wheel in an overarching synthesis of systematic reviews. BMJ Open 2019 Jun;9(6):e024950. [doi: 10.1136/bmjopen-2018-024950]
19. Buchanan H, Newton JT, Baker SR, Asimakopoulou K. Adopting the COM-B model and TDF framework in oral and dental research: a narrative review. Community Dent Oral Epidemiol 2021 Oct;49(5):385-393. [doi: 10.1111/cdoe.12677] [Medline: 34270110]
20. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [Medline: 10940958]
21. Boddy CR. Sample size for qualitative research. QMR 2016 Sep 12;19(4):426-432. [doi: 10.1108/QMR-06-2016-0053]

22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Wong AKC, Wong FKY. The psychological impact of a nurse-led proactive self-care program on independent, non-frail community-dwelling older adults: a randomized controlled trial. Int J Nurs Stud 2020 Oct;110:103724. [doi: 10.1016/j.ijnurstu.2020.103724] [Medline: 32777605]
24. Wong AKC, Wong FKY, Chow KKS, Wong SM, Bayuo J, Ho AKY. Effect of a mobile health application with nurse support on quality of life among community-dwelling older adults in Hong Kong: a randomized clinical trial. JAMA Netw Open 2022 Nov 1;5(11):e2241137. [doi: 10.1001/jamanetworkopen.2022.41137] [Medline: 36350651]
25. Wong AKC, Bayuo J, Su JJ, et al. Exploring the experiences of community-dwelling older adults on using wearable monitoring devices with regular support from community health workers, nurses, and social workers: qualitative descriptive study. J Med Internet Res 2024 Aug 7;26:e49403. [doi: 10.2196/49403] [Medline: 39110493]

Abbreviations

AI: artificial intelligence
COM-B: Capability, Opportunity, Motivation, and Behavior
NLP: natural language processing
TDF: Theoretical Domains Framework

Edited by F Puga; submitted 23.09.24; peer-reviewed by HJ Kim, L Zhou, A Nakagomi; revised version received 23.12.24; accepted 07.01.25; published 12.02.25.

Please cite as:
Wong AKC, Lee JHT, Zhao Y, Lu Q, Yang S, Hui VCC
Exploring Older Adults' Perspectives and Acceptance of AI-Driven Health Technologies: Qualitative Study
JMIR Aging 2025;8:e66778
URL: https://aging.jmir.org/2025/1/e66778
doi:10.2196/66778

Original Paper

Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study

Shayan Nejadshamsi[1,2,3], PhD; Vania Karami[1,2,3], PhD; Negar Ghourchian[4], PhD; Narges Armanfard[1,5], PhD; Howard Bergman[2], MD, FRCP; Roland Grad[2], MD, CM, MSc; Machelle Wilchesky[2,3,6], PhD; Vladimir Khanassov[2], MD; Isabelle Vedel[2], MD, PhD; Samira Abbasgholizadeh Rahimi[1,2,3,7], BEng, PhD

[1]Mila—Quebec Artificial Intelligence Institute, Montreal, QC, Canada

[2]Family Medicine Department, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC, Canada

[3]Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada

[4]Aerial Technologies, Montreal, QC, Canada

[5]Department of Electrical and Computer Engineering, Faculty of Engineering, McGill University, Montreal, QC, Canada

[6]Donald Berman Maimonides Centre for Research in Aging, Montreal, QC, Canada

[7]Faculty of Dental Medicine and Oral Health Sciences, McGill University, Montreal, Canada

Corresponding Author:
Samira Abbasgholizadeh Rahimi, BEng, PhD
Family Medicine Department
Faculty of Medicine and Health Sciences
McGill University
5858 Côte des Negies
Montreal, QC, H3S 1Z1
Canada
Phone: 1 5143987375
Email: samira.rahimi@mcgill.ca

Abstract

Background: Depression, characterized by persistent sadness and loss of interest in daily activities, greatly reduces quality of life. Early detection is vital for effective treatment and intervention. While many studies use wearable devices to classify depression based on physical activity, these often rely on intrusive methods. Additionally, most depression classification studies involve large participant groups and use single-stage classifiers without explainability.

Objective: This study aims to assess the feasibility of classifying depression using nonintrusive Wi-Fi–based motion sensor data using a novel machine learning model on a limited number of participants. We also conduct an explainability analysis to interpret the model's predictions and identify key features associated with depression classification.

Methods: In this study, we recruited adults aged 65 years and older through web-based and in-person methods, supported by a McGill University health care facility directory. Participants provided consent, and we collected 6 months of activity and sleep data via nonintrusive Wi-Fi–based sensors, along with Edmonton Frailty Scale and Geriatric Depression Scale data. For depression classification, we proposed a HOPE (Home-Based Older Adults' Depression Prediction) machine learning model with feature selection, dimensionality reduction, and classification stages, evaluating various model combinations using accuracy, sensitivity, precision, and F_1-score. Shapely addictive explanations and local interpretable model-agnostic explanations were used to explain the model's predictions.

Results: A total of 6 participants were enrolled in this study; however, 2 participants withdrew later due to internet connectivity issues. Among the 4 remaining participants, 3 participants were classified as not having depression, while 1 participant was identified as having depression. The most accurate classification model, which combined sequential forward selection for feature selection, principal component analysis for dimensionality reduction, and a decision tree for classification, achieved an accuracy of 87.5%, sensitivity of 90%, and precision of 88.3%, effectively distinguishing individuals with and those without depression. The explainability analysis revealed that the most influential features in depression classification, in order of importance, were

"average sleep duration," "total number of sleep interruptions," "percentage of nights with sleep interruptions," "average duration of sleep interruptions," and "Edmonton Frailty Scale."

Conclusions: The findings from this preliminary study demonstrate the feasibility of using Wi-Fi–based motion sensors for depression classification and highlight the effectiveness of our proposed HOPE machine learning model, even with a small sample size. These results suggest the potential for further research with a larger cohort for more comprehensive validation. Additionally, the nonintrusive data collection method and model architecture proposed in this study offer promising applications in remote health monitoring, particularly for older adults who may face challenges in using wearable devices. Furthermore, the importance of sleep patterns identified in our explainability analysis aligns with findings from previous research, emphasizing the need for more in-depth studies on the role of sleep in mental health, as suggested in the explainable machine learning study.

(JMIR Aging 2025;8:e67715) doi:10.2196/67715

KEYWORDS

depression; classification; machine learning; artificial intelligence; older adults

Introduction

Depression is a prevalent mental health disorder characterized by emotional dysregulation, leading to persistent sadness, loss of interest, and anhedonia [1-3]. The rising incidence of depression among older adults has become a significant public health issue [4-6]. Early detection of depression and corresponding intervention are vital for improving mental health outcomes and reducing the overall burden on individuals and health care systems [7-9]. Traditional methods for assessing depression include various approaches that typically require in-person evaluations, specialized training in comprehensive geriatric assessments, and reliance on clinical judgment and questionnaires, which can be challenging and resource-intensive [10-12]. These methods require older adults to visit clinical settings frequently, increasing strain on health care facilities and reducing data collection opportunities. Additionally, many older adults prefer to remain in their homes and be remotely monitored in that environment, highlighting the need for remote care solutions in this demographic [13,14].

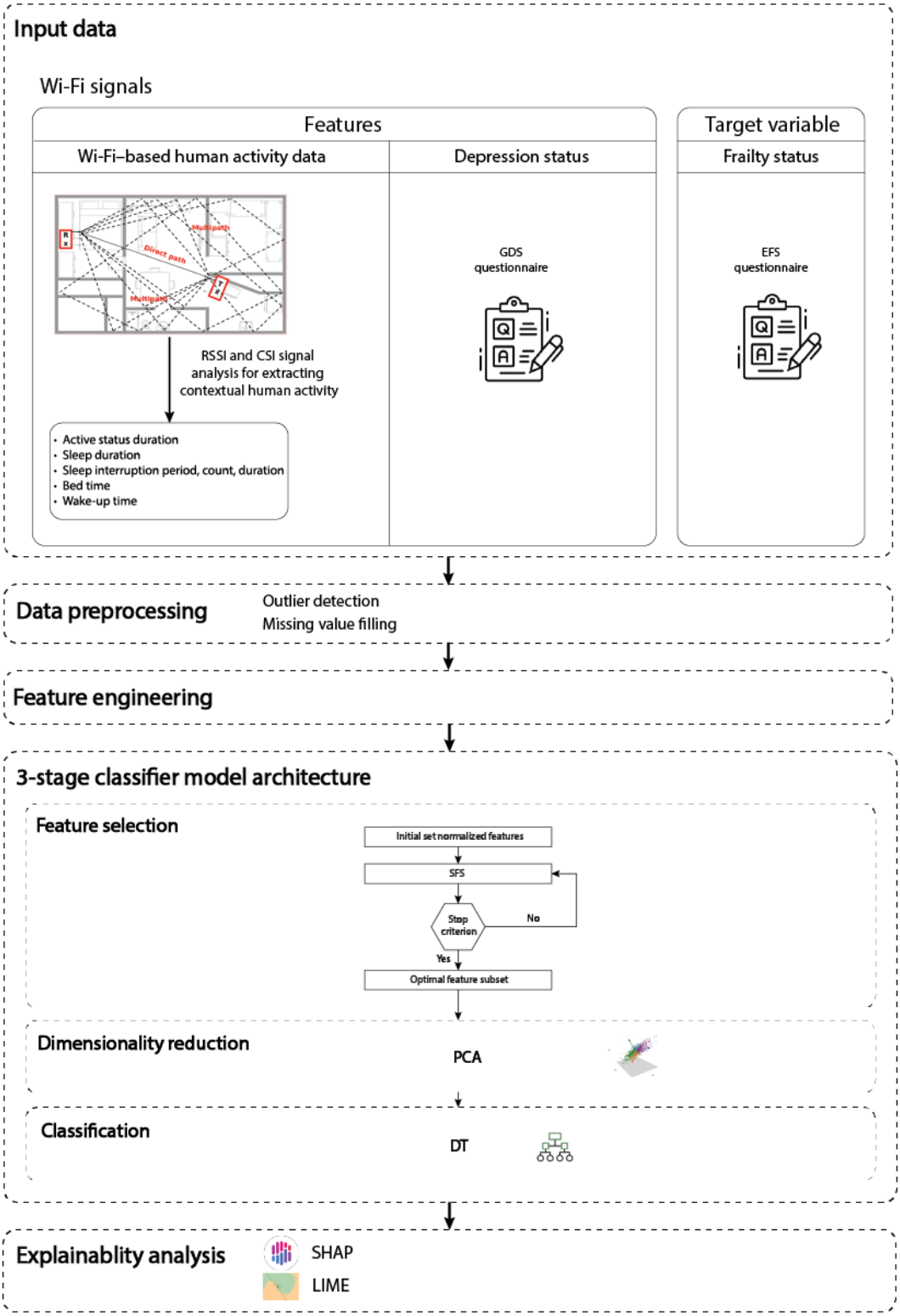
Physical activity and mobility are among the important factors in evaluating depression, with strong correlations established between these parameters and depression assessments [1,15]. The advent of the Internet of Things has enabled continuous and remote monitoring of physical activity. Several studies have used statistical methods to analyze the relationship between physical activity, as measured by wearable devices, and depression [16-21]. As the field of artificial intelligence (AI) advances, machine learning models have emerged as promising tools for depression classification using physical activity data [22]. For instance, Adamczyk and Malawski [23] used data from wearable actigraph watches in 3 classification models: logistic regression (LR), support vector machine (SVM), and random forest (RF) comparing automatic and manual feature engineering for depression classification. Bai et al [24] used phone use, sleep data, and step counts from 334 participants, using 2 feature selection methods (L1-based feature selection) and 6 machine learning models (decision tree [DT], k-nearest neighbors, naive Bayes, LR, SVM, and RF) for mood classification. Chikersal et al [25] analyzed data from smartphones and fitness trackers of 138 college students to identify those experiencing depressive symptoms, using nested randomized LR for feature selection and AdaBoost with gradient boosting classifier. Dai et al [26] used heart rate, energy expenditure, sleep, and other activity data from wearable Fitbit devices for depression remission detection in 106 participants within 2 intervention and control groups, using a multitask learning algorithm comprising 2 dense layers with shared parameters. Similarly, Griffiths et al [27] classified depression using activity and sleep data from Fitbit devices of 24 participants through an RF model. Espino-Salinas et al [28] used wrist-worn accelerometers to measure physical activity in 55 participants, applying a 2D-convolutional neural network (CNN) and a deep neural network for depression classification. Jakobsen et al [29] used RF, deep neural network, and CNN algorithms for depression classification with wrist-worn actigraph data from 55 participants. Jung et al [30] used gait accelerometry data and a bidirectional long short-term memory network–based classifier to assess depression in 45 older adults. Other studies have also explored the use of wearable device data combined with various classification methods, such as 1D-CNN [31], deep convolutional neuro-fuzzy [32], Ensemble models [33], and extreme gradient boosting [34].

These recent studies highlight the integration of wearable devices with machine learning algorithms as a promising approach for continuous, remote, home-based monitoring and early detection of depression. However, wearable device-based approaches face challenges due to their intrusive nature [35]. Participants are required to wear sensors or devices, which may lead to issues with compliance, comfort, and data accuracy, especially over extended periods [36], particularly among older adults [37]. These challenges can result in inaccurate data collection [38]. Nonwearable methods, such as remote monitoring through ambient sensors, offer potential solutions to these issues [37,39,40]. These approaches can alleviate concerns related to device adherence and physical discomfort, providing a more seamless integration into daily life. Additionally, our literature review indicates that most studies on depression classification use a relatively large number of participants and primarily use single-stage classifiers. Many studies also focus on detailed aspects of physical activity data (eg, body displacement, acceleration) using intrusive wearable sensors, which, while effective, present challenges related to user comfort and compliance. This reliance on wearable technology underscores the need for exploring alternative, less intrusive methods.

Recently, Wi-Fi–based sensing in smart homes has emerged as an alternative method for detecting and monitoring contextual human activity and movement [41]. Wi-Fi–based technologies are increasingly adopted due to their existing infrastructure in homes and minimal additional setup costs [42]. These technologies use signal metrics such as received signal strength indicator (RSSI) and channel state information (CSI) to analyze Wi-Fi signal characteristics, offering human activity identification compared to invasive wearable sensors, image analysis, or video-based systems [43]. Leveraging Wi-Fi–based activity data with machine learning presents a viable approach for different diseases, specifically depression classification, which is the focus of this study.

Technological and AI-based methods for depression classification predominantly rely on wearable sensors and are often conducted on large participant groups and typically use statistical techniques or single-stage machine learning classifiers. Furthermore, these approaches often overlook the explainability analysis of the models, which causes a lack of understanding of the underlying decision-making processes and the contribution of individual features to the model's predictions. To address these challenges, this feasibility study introduces a novel 3-stage machine learning model that incorporates feature selection, dimensionality reduction, and classification for depression classification. This model is specifically designed for a limited number of participants using low cost, easily installable Wi-Fi data, which provides continuous insights into human indoor activities. In addition to creating a model that functions effectively with small sample sizes, this study integrates explainable AI techniques to enhance the interpretability of the model's predictions. This approach ensures that the insights derived from the model are transparent and comprehensible, providing clarity on how specific features contribute to the classification outcomes. Figure 1 illustrates a schematic overview of our proposed framework.

Figure 1. Structure of the automatic Wi-Fi–based depression classification framework. CSI: channel state information; DT: decision tree; EFS: Edmonton Frailty Scale; GDS: Geriatric Depression Scale; LIME: local interpretable model-agnostic explanations; PCA: principal component analysis; RSSI: received signal strength indicator; SFS: sequential forward selection; SHAP: Shapely addictive explanations.



Methods

Data Acquisition

We begin by outlining the data acquisition process for the participants, followed by an in-depth explanation of our proposed HOPE model.

Study Cohort

Our recruitment approach used both digital and in-person strategies. Digital outreach was conducted through email campaigns, social media platforms, and digital posters. Prospective participants were provided with detailed information about the study, and those who expressed interest received a consent form. A member of the research team then coordinated

the setup of the monitoring equipment. Participants were compensated for their time and involvement with an e-gift card.

Participants in this study were required to be aged 65 years or older, capable of communicating in English or French, and have access to an internet connection at home. Exclusion criteria were as follows: (1) individuals with mental or physical conditions that would impede their ability to participate in the study and its 6-month follow-up, such as gait or balance disorders, active mental health issues, or the use of mobility aids such as canes; and (2) individuals with current substance use disorder, including alcohol or drugs, due to their potential impact on physical mobility. However, individuals with a history of substance use disorder who were no longer consuming were considered eligible.

Experimental Protocol

This study used a nonintrusive Wi-Fi–based motion sensor system to facilitate remote monitoring of human activity. Initially, we collected demographic information (ie, age and gender) from the participants. Subsequently, Wi-Fi data were acquired through our remote monitoring technology [44]. This device, installed at the network access point, detects Wi-Fi signals in environments conducive to passive sensing, including private residences and public spaces where Wi-Fi is prevalent [45]. In such indoor settings, Wi-Fi signals exhibit stability in the absence of individuals but fluctuate significantly with the presence and movement of people [45]. These signal variations correspond to distinct patterns associated with human movements and activities, thus providing valuable data for activity monitoring [46]. The collected data were subsequently transferred to secure cloud storage via an internet connection. Our team has developed advanced signal processing and AI-based algorithms to process raw Wi-Fi RSSI and CSI measurements. These algorithms standardize signal variations and translate RSSI and CSI fluctuations into a detailed set of contextual information related to human activity [45]. This information encompasses daily activity duration, bedtime, wake-up time, total sleep duration, and sleep interruption information. In addition to using Wi-Fi–based activity data, frailty and depression statuses were also assessed using validated assessment tools. The Edmonton Frailty Scale (EFS), which measures multiple dimensions of frailty [47-49], and the 15-item Geriatric Depression Scale (GDS) [50] were administered at the end of the experiment.

Analytical Framework

The methodical steps of our proposed model are detailed in the following subsections. During the development, implementation, and reporting, we adhered to the Minimum Information About Clinical Artificial Intelligence Modeling (MI-CLAIM) guidelines [51], following best practices designed to promote transparency and reproducibility of our AI model.

Data Preparation and Feature Extraction

To prepare the contextual human activity data as input for depression classification, obtained from our Wi-Fi signal analysis software, we designed a preprocessing stage. This process involves handling missing values and outliers to ensure the integrity of the data [52,53]. To enhance analytical depth and improve model performance, we implemented feature engineering on contextual human activity data. This process involves extracting a variety of new features, such as the mean and SD of bedtime and wake-up times, mean and SD of sleep duration (in hours), total count and mean of sleep interruptions, total duration and mean duration of sleep interruptions (in hours), longest continuous sleep duration (in hours), percentage of nights with sleep disturbances, and metrics related to daily activity, including the mean and SD of total daily activity and hourly activity durations, as well as peak activity hour. These derived features, combined with EFS data, were incorporated into our 3-stage machine learning classification model. The depression status, determined using the GDS, was used to label the samples for classification purposes.

HOPE Model Development

The proposed HOPE Model was designed for depression classification in older adults using nonintrusive Wi-Fi–based motion sensor data. Due to the limited number of participants, it is necessary to provide an efficient pipeline for preprocessing, feature extraction, and classification. The limited number of participants and high dimensionality of features required a tailored multistage machine learning pipeline to maximize classification accuracy. Furthermore, ensuring that our depression classification model is explainable to clinicians is important, as highlighted in our previous works [54,55]. To address this, we incorporated explainable machine learning techniques such as Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME). To achieve these goals, our proposed HOPE model was structured into 3 stages of machine learning architecture: feature selection, dimensionality reduction, and classification followed by post hoc explainability analysis using SHAP and LIME. Each stage plays a critical role in refining the data and ensuring that the final classification is both accurate and interpretable. Table 1 provides details on the various techniques used and evaluated at each stage.

Table 1. Methods used at each phase of our 3-stage architecture.

Feature selection	Dimensionality reduction	Classification	Explainability analysis
CFS[a] [56]	PCA[b] [57]	NB[c] [58]	SHAP[d] [59]
SFS[e] [60]	FA[f] [61]	LR[g] [62]	LIME[h] [63]
MI[i] [64]	LDA[j] [65]	kNN[k] [66]	—[l]
SelectKBest [67]	kPCA[m] [68]	SVM[n] [69]	—
RFE[o] [70]	—	Decision tree [71]	—
—	—	RF[p] [72]	—
—	—	GBM[q] [73]	—
—	—	XGboost [74]	—
—	—	LightGBM [75]	—
—	—	Voting classifier [76]	—
—	—	Bagging classifier [77]	—
—	—	AdaBoost [78]	—

[a]CFS: correlation-based selection.
[b]PCA: principal component analysis.
[c]NB: naive Bayes.
[d]SHAP: Shapley addictive explanations.
[e]SFS: sequential forward selection.
[f]FA: factor analysis.
[g]LR: logistic regression.
[h]LIME: local interpretable model-agnostic explanations.
[i]MI: mutual information.
[j]LDA: linear discriminant analysis.
[k]kNN: k-nearest neighbor.
[l]Not applicable.
[m]kPCA: kernel principal component analysis.
[n]SVM: support vector machine.
[o]RFE: recursive feature elimination.
[p]RF: random forest.
[q]GBM: gradient boosting machine.

Feature selection is performed to reduce the dimensionality of the dataset by identifying the most relevant features for depression classification, enhancing both the speed and accuracy of the classification model [79]. The reduced subset of features serves as input for the subsequent dimensionality reduction stage. The validity of the chosen selected features is investigated using correlation analysis in the Results section. Dimensionality reduction techniques are applied to further refine the feature set compared to the initial feature selection stage [80] and to minimize overfitting. The dimensionally reduced features from the second stage were then processed in the third stage, which focused on classification. In this stage, the classification model processes the features derived from the earlier stages to categorize samples into 2 target classes: "participants with depression" and "participants without depression." The classification algorithm leverages the patterns identified in the features, such as sleep duration and interruptions, to make predictions. The classification task involved assigning a probability score to each sample, determining the likelihood of belonging to either class based on the relationships in the feature set. A decision boundary was then established to assign the final class label for each sample. The classification process was systematically evaluated to ensure robustness and reliability, focusing on separating the 2 groups effectively even with the small dataset. The machine learning classification pipeline was designed to minimize the risk of overfitting by using techniques such as feature selection and dimensionality reduction, ensuring that the most informative and relevant features were used for prediction. To conclude, we used explainable AI techniques to interpret the model's predictions, focusing on identifying the most influential features and their impact on classification outcomes.

Following the training and evaluation of all potential combinations for each stage, the architecture using sequential forward selection (SFS) for feature selection, principal component analysis (PCA) for dimensionality reduction, and DT for classification emerged as the most effective configuration. This SFS-PCA-DT framework (Figure 1)

demonstrated superior performance compared to other combinations.

Our proposed model is supported by different considerations. The SFS algorithm incrementally selects features to improve classification performance and is particularly suited for datasets with a smaller number of participants [81]. Unlike filter methods, SFS acts as a wrapper technique that pairs with a machine learning classification algorithm, providing greater stability in performance [81]. The method starts with no selected features and progressively adds them based on their ability to enhance cross-validation outcomes. In the second stage, PCA serves as a highly effective tool for reducing the dimensionality of data in an unsupervised manner [82]. It converts the initial set of features into a reduced number of uncorrelated components, maintaining the bulk of the data's variance. This reduction is important for preventing overfitting, especially with limited sample sizes. As a classification algorithm, DT algorithm, known for its strength in binary classification tasks, effectively uses the streamlined feature set generated in earlier stages. It models data by learning simple decision rules inferred from the input features, creating a tree-like structure. Each internal node in the tree represents a decision based on a feature, each branch represents an outcome of the decision, and each leaf node represents a class label [83]. The integration of SFS, PCA, and DT results in an efficient model that aligns with established methodologies and theoretical principles in the field. In this study, comprising only 4 participants, the training and validation procedure were carefully designed to minimize overfitting and to achieve reliable model generalization. To this end, we used a 4-fold cross-validation strategy. Each fold consisted of 3 participants for training and 1 participant for testing, ensuring that every participant contributed to both training and testing in separate iterations. This approach was repeated 10 times with different random seeds to account for variations in the training process, further enhancing the robustness of the performance metrics. During the training phase, a range of hyperparameter optimization techniques, including random search [84], Bayesian optimization [85], and Hyperband [86] was performed for each component of the 3-stage pipeline. For example, the DT classifier's maximum depth, minimum sample split, and criterion parameters are tuned using Bayesian optimization within predefined search spaces. Similarly, for PCA, the optimal number of components is optimized to maximize variance retention while preventing overfitting. The SFS algorithm was guided by internal cross-validation within the training set to identify the most predictive subset of features.

Evaluation Metrics

To validate the effectiveness of our depression classification method, we used 4 evaluation metrics: accuracy [87], sensitivity [87], precision [87], and F_1-score. Accuracy provides a comprehensive measure of the model's overall performance. Sensitivity helps ensure that the model accurately identifies as many true cases of depression as possible, minimizing the risk of missing individuals who actually have the condition [88]. To further validate the stability of the model, we present the training and test accuracies against the hyperparameter variations, demonstrating the model convergence.

Ethical Considerations

The study received approval from McGill University's Institutional Ethics Committee (A06-B18-21A), allowing data collection and analysis for this project. Written informed consent was collected from all participants prior to their involvement in the study. All collected data were anonymized immediately after collection, with no personally identifiable information retained to ensure participant confidentiality. Participants received a $20 e-gift card as compensation for their time. At every stage of the research, we adhered to the ethical principles outlined in the Declaration of Helsinki [89] and the Tri-Council Policy Statement [90].

Results

Clinical Study Insights

Six community-dwelling older adults residing in Montreal, Canada, were recruited for this study between May 2022 and September 2022. However, 2 participants withdrew due to internet connectivity issues, resulting in the use of data from the remaining 4 participants for the analysis. The EFS results indicated that 2 participants exhibited moderate frailty (scores ranging from 6 to 11), while the other 2 participants were classified as nonfrail (scores of 5 or below). GDS results suggested that 1 participant had depressive symptoms (score=10), while the other 3 participants did not (score range=0-4).

Some of the participants' demographic and clinical characteristics are detailed in Table 2. Over a 6-month period, the activities of each participant were continuously monitored at 15-minute intervals using Wi-Fi motion sensors. Following the identification of potential input features derived from Wi-Fi signals and questionnaire data, we designed and developed a 3-stage architecture as outlined in the methodology section. To support future research and ensure the reproducibility of our findings, the code for our model is openly accessible on our lab's GitHub repository [91].

Table 2. Demographic and clinical characteristics of the participants of this study.

	Participant without depression (n=3)	Participant with depression (n=1)
Sex (female/male), n/n	1/2	1/0
Age (years), mean (SD)	67.05 (3.70)	65.50 (0.00)
Edmonton frailty scale, mean (SD)	3.34 (3.21)	7.00 (0.00)
Geriatric depression scale, mean (SD)	4.00 (1.00)	10.00 (0.00)

We assessed a total of 240 model configurations, which resulted from the combination of 5 feature selection methods, 4 dimensionality reduction techniques, and 12 classifiers. Each configuration was trained 10 times to account for variations in performance metrics, ensuring the robustness of our findings. To mitigate the overfitting risk, we used k-fold cross-validation [92]. We experimented with different initial feature sets to optimize model performance, and among the various hyperparameter tuning methods, Bayesian optimization consistently yielded superior results.

Model Performance and Validation

This section presents the outcomes of our classification model evaluation. Table 3 summarizes the top-performing configurations among the 240 model variations that we tested. The SFS-PCA-DT model, which integrates SFS, PCA, and DT, emerged as the leading performer across multiple metrics. Its relatively high accuracy indicates the model's ability to distinguish between individuals with and without depression. The model's high sensitivity ensures that individuals with depression are correctly identified. This is critical in clinical settings, where depression, particularly among older adults, often goes unrecognized despite its severe impact on cognitive function [93-96], quality of life [97,98], and mortality risk [99,100]. Ensuring high sensitivity reduces the likelihood of missed diagnoses, which is important for timely and effective treatment. However, the relatively high standard deviation in both accuracy and sensitivity suggests that the model's performance may vary, indicating occasional instances of less reliable predictions.

Table 3. Average classification performance of the top 5 architectures.

Model	Accuracy (%)	Sensitivity (%)	Precision (%)	F_1-score (%)
SFS[a] – PCA[b] – DT[c]	87.50 (12.50)	90.00 (20.00)	88.34 (18.34)	86.00 (14.74)
SFS + FA[d] + DT	85.00 (16.58)	83.34 (25.82)	90.00 (20.00)	81.67 (18.93)
SFS + PCA + LR[e]	82.50 (19.53)	85.00 (22.91)	83.34 (25.82)	80.00 (20.82)
SFS + PCA + SVM[f]	82.50 (19.53)	90.00 (20.00)	75.00 (22.42)	80.00 (20.82)
MI[g] + PCA + LR	80.00 (10.00)	85.00 (22.91)	78.34 (22.42)	76.00 (13.06)

[a]SFS: sequential forward selection.

[b]PCA: principal component analysis.

[c]DT: decision tree.

[d]FA: factor analysis.

[e]LR: logistic regression.

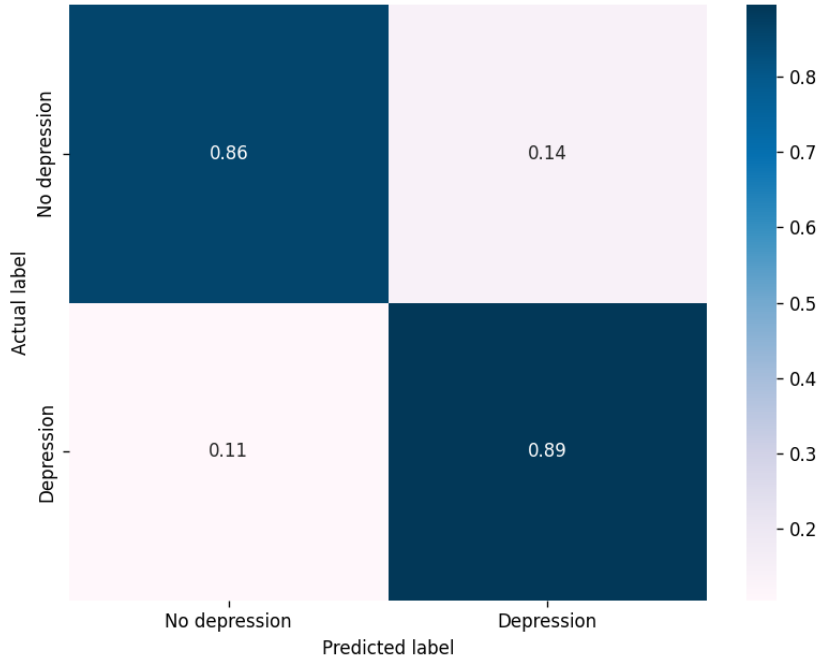
[f]SVM: support vector machine.

[g]MI: mutual information.

Table 3 illustrates that SFS is frequently featured among the highest-performing models, including the top model with 87.50% accuracy, and 3 other strong contenders. Mutual information also proves effective, appearing in 1 model with 80.00% accuracy, indicating that both SFS and mutual information are potent feature selection techniques for depression classification with limited samples. PCA seems to be the preferred method for dimensionality reduction, being used in 4 of the top 5 models. Among classifiers, DT stands out, featuring in the top 2 models with 87.50% and 85.00% accuracy. Other classifiers, such as LR and SVM, also perform well, each appearing in models with an accuracy exceeding 80%. The results highlight interesting tradeoffs, such as the SFS + PCA + SVM model, which, while slightly lower in accuracy (82.50%), maintains a high sensitivity (90.00%). This supports the practice of evaluating models using multiple metrics, especially in situations where the application involves clinical diagnoses, where accurately identifying true positives is crucial.

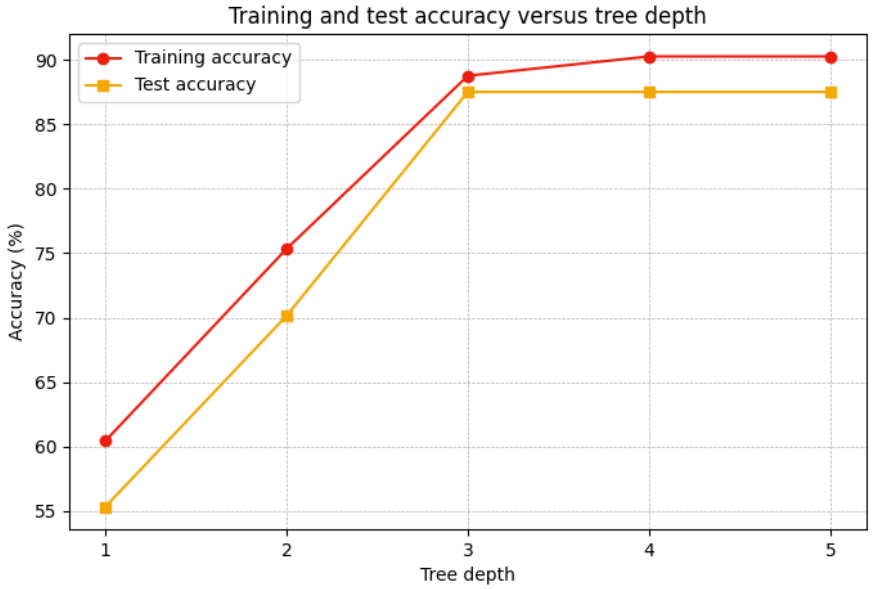
Figure 2 displays the averaged confusion matrix for the top-performing SFS – PCA – DT model used to classify depression status. Due to the constraint of having only 4 samples, we used a 4-fold cross-validation strategy, with each fold being tested 10 times to ensure a thorough evaluation. The model showed promising results, accurately classifying 86% of individuals without depression as not having depression and 89% of individuals with depression as having depression. The average false positive rate, where individuals without depression were incorrectly classified as having it, was 14%, while the average false negative rate, where individuals with depression were incorrectly classified as not having it, was 11%.

Figure 2. Confusion matrix for the top-performing model (SFS – PCA – DT). DT: decision tree; PCA: principal component analysis; SFS: sequential forward selection.



To validate the convergence of the proposed algorithm, we analyzed the relationship between tree depth and accuracy on both the training and test datasets. Figure 3 demonstrates the training and test accuracy of the best performing model (SFS – PCA – DT) as a function of tree depth. The training accuracy increases consistently with tree depth, stabilizing at its maximum, reflecting that the model can fully capture the training data as depth increases. The test accuracy improves initially with increasing tree depth but stabilizes beyond a depth of 3. These observations confirm that the proposed algorithm achieves convergence in terms of performance tradeoffs between model complexity and generalization.

Figure 3. Training and test accuracy as a function of tree depth, demonstrating convergence of our proposed model.

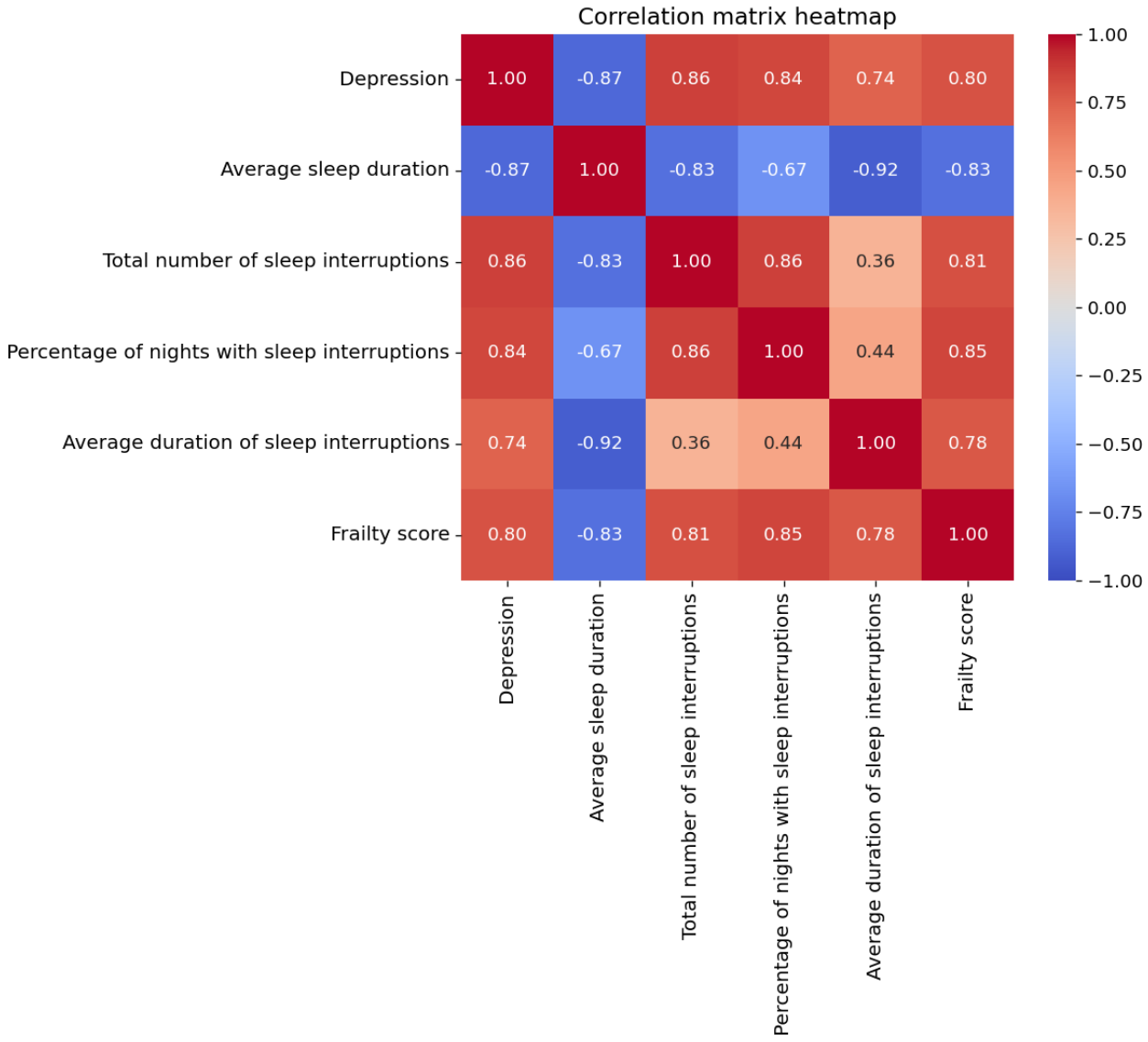


Feature Selection and Analysis

In our 3-stage classification model, we implement a combination of feature selection and dimensionality reduction techniques to improve the efficacy of our machine learning approach [80]. The features selected by SFS included “average sleep duration,” “total number of sleep interruptions,” “percentage of nights with sleep interruptions,” “average duration of sleep interruptions,” and “EFS.” Correlation analysis of these features revealed notable associations with depression status. Such analysis helps identify how variations in these features might be related to changes in depression, providing valuable insights for clinicians and researchers to develop more effective diagnostic tools and treatments.

Figure 4 illustrates a strong negative correlation between depression and average sleep duration. Conversely, depression was positively correlated with the total number of sleep interruptions, percentage of nights with sleep interruptions, average duration of sleep interruptions, and EFS. Blue cells indicate negative correlation values, while red cells represent positive correlations. Darker colors signify stronger correlations. The high correlation values highlight the significance of these factors in understanding and potentially classifying older adults with depression, aligning with findings from previous studies [101-105]. For example, Vallance et al [106] demonstrated that engaging in daily activities can alleviate the adverse effects of depression among older adults. Furthermore, several studies have highlighted a connection between depression and frailty [107,108]. Vaughan et al [109] showed that the prevalence of both depression and frailty among individuals aged 55 years and older exceeds 10%. These findings confirm the association between depression and the features incorporated in our model.

Figure 4. Correlation matrix heatmap between depression and selected features by sequential forward selection.



Comparative Analysis With Baseline Models

In this section, we evaluate the performance of our proposed 3-stage architecture for depression classification, which leverages Wi-Fi–based contextual human activity data, against baseline models previously outlined in the introduction. Although direct comparisons are inherently difficult due to differences in data acquisition methods (wearable devices), feature sets, and sample sizes, this analysis serves to contextualize the effectiveness of our approach for our case study with a limited number of participants.

For the baseline models, we selected the most current classification architectures used in the literature for depression classification. To ensure a fair comparison, each baseline model was trained and tested using the same feature set applied in our experiment. The resulting performance metrics for each baseline model are presented in Table 4. Despite the challenges associated with our smaller sample size, the comparison offers valuable insights into the relative efficacy of our method.

Table 4. Average performance across different baseline machine learning models.

Model architecture	Accuracy (%)	Sensitivity (%)	Precision (%)	F_1-score (%)
RF[a] [23,27]	12.50	5.00	5.00	N/A[b]
SVM[c] [23]	15.00	10.00	10.00	N/A
LR[d] [23]	22.50	15.00	13.34	N/A
XGBoost [34]	25.00	10.00	10.00	N/A
L1-based feature selection + DT[e] [24]	32.50	35.00	18.34	N/A
L1-based feature selection + RF [24]	22.50	15.00	13.34	N/A
L1-based feature selection + kNN[f] [24]	22.50	15.00	13.34	N/A
L1-based feature selection + NB[g] [24]	30.00	25.00	15.00	N/A
L1-based feature selection + LR [24]	37.50	45.00	25.00	N/A
L1-based feature selection + SVM [24]	25.00	10.00	10.00	N/A
Randomized LR + AdaBoost [25]	55.00	73.33	55.00	N/A
HOPE model[h]	87.50	90.00	88.34	86.00

[a]RF: random forest.

[b]N/A: data not applicable.

[c]SVM: support vector machine.

[d]LR: logistic regression.

[e]DT: decision tree.

[f]kNN: k-nearest neighbor.

[g]NB: naive Bayes.

[h]Best performed proposed model.

As shown in Table 4, traditional single-stage machine learning classifiers such as RF, LR, and SVM demonstrate relatively lower performance, with accuracy ranging from 12.50% to 22.50%. Among these, LR achieves relatively higher accuracy. XGBoost exhibits better performance than the traditional models. Incorporating feature selection techniques further improves the performance of these models. Specifically, combining L1-based feature selection with various classifiers results in modest performance gains, while the randomized LR combined with AdaBoost achieves a significant improvement, reaching 55.00% accuracy and 73.33% sensitivity. Our proposed 3-stage architectures significantly surpass all other baseline models across all metrics. While most baseline models struggle with sensitivity and precision, often scoring below 15%, our best proposed model demonstrates substantial enhancements in these metrics with 90.00% sensitivity and 88.34% precision, indicating a superior capability to correctly identify positive cases and reduce false positives.

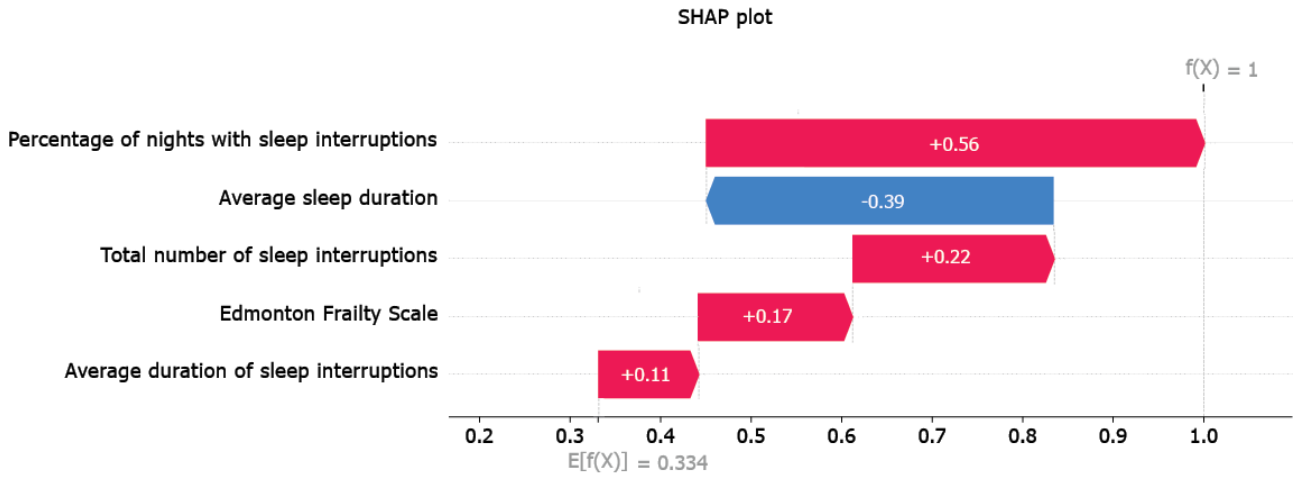
Model Explainability

To enhance our understanding of the decision-making processes within our proposed model, we used SHAP [50] and LIME [54] for model interpretability analysis. These model-agnostic methods can be applied across various machine learning models, providing valuable insights into our model's predictive behavior. By integrating these interpretability techniques, we aim to improve the transparency and potential clinical relevance of our depression classification framework. These methods help us identify which features most significantly influence the model's predictions, particularly in the context of depression classification.

SHAP Analysis

The SHAP waterfall plot (Figure 5) illustrates the relative importance of features for depression classification, with red and blue colors representing positive and negative contributions, respectively. Among the features, "the percentage of nights with sleep interruptions" is the most impactful, positively correlating with depression risk, indicating that frequent sleep disturbances are a strong predictor of depression. Conversely, the average sleep duration exhibits a substantial negative impact on depression prediction, suggesting that longer sleep durations are associated with a reduced likelihood of depression. Sleep-related variables continue to play a pivotal role in the model's predictions; both the total number of sleep interruptions and the average duration of these interruptions contribute positively to depression risk, further underscoring the importance of uninterrupted sleep in depression diagnosis. Although the frailty scale is included in the model, its influence is relatively minor compared to sleep-related features.

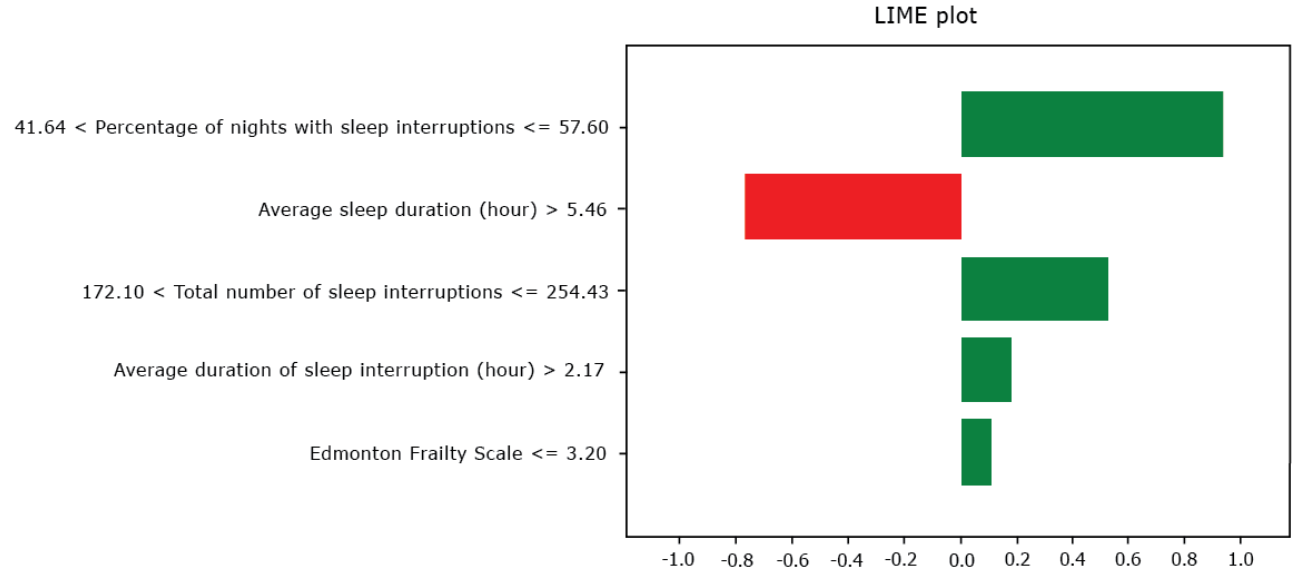
Figure 5. SHAP analysis. SHAP: Shapley additive explanations.



LIME Analysis

The LIME plot (Figure 6) provides a complementary view of feature importance, with green and red colors indicating positive and negative influences, respectively. Consistent with the SHAP results, LIME identifies "the percentage of nights with sleep interruptions" as the most critical feature in the classification of depression of our proposed model. Similarly, the average sleep duration is shown to have a significant negative impact on depression classification, in line with SHAP findings. The total number of sleep interruptions also ranks highly with a positive influence on depression risk, again aligning with SHAP results. A notable difference between the 2 methods is the relatively lower impact of the frailty scale in the LIME analysis, which requires further investigations.

Figure 6. LIME analysis. LIME: local interpretable model-agnostic explanations.



Discussion

Our research uses Wi-Fi–based motion sensors to extract daily activities, which are then used in our proposed machine learning method for depression classification.

Our study findings confirm the feasibility of using Wi-Fi–based motion sensors for depression classification among older adults. Our proposed HOPE (Home-Based Older Adults' Depression Prediction) model achieved an accuracy of 87.5%, sensitivity of 90%, and precision of 88.3%. The most influential features identified were sleep-related metrics, such as average sleep duration and sleep interruptions, highlighting the importance of sleep patterns in depression classification. These findings suggest that Wi-Fi–based monitoring offers a nonintrusive and effective alternative to conventional wearable technologies for depression assessment. These conventional methods, while effective, often present challenges in terms of participant compliance, particularly among older adults, due to their burdensome and sometimes uncomfortable nature. In contrast, our Wi-Fi–based approach is nonintrusive and allows for continuous monitoring without requiring participants to wear or interact with any devices. This can significantly enhance participant compliance and the integrity of the data collected over extended periods. Compared to other nonintrusive monitoring technologies, such as camera-based methods, our Wi-Fi–based approach has distinct advantages. Wi-Fi infrastructure is prevalent in most homes and does not pose

privacy risks, making it a cost-effective and scalable solution for continuous health monitoring. Furthermore, unlike previous studies that rely on microlevel body displacement and accelerometer data, our study emphasizes macrolevel physical activity features like sleep patterns and overall activity levels shift is crucial as it highlights the potential of using broader, more easily obtainable metrics to assess depression status. Our findings demonstrate that these macrolevel features are not only feasible but also effective measures for depression classification, broadening the scope of nonintrusive monitoring technologies in mental health research. The next steps can be extracting more detailed types of human activity using nonintrusive Wi-Fi data and expanding more on using this type of data acquisition for depression classification.

Additionally, our proposed model demonstrates relatively high performance compared to other classification models presented in existing depression classification studies, even with a limited sample size. Many studies using physical activity data from wearable devices often benefit from larger datasets and frequently use single-stage classifiers or deep neural networks. These models generally show strong performance with abundant data; however, their effectiveness diminishes when applied to smaller datasets, such as the one in this feasibility study. To address the limitations imposed by our smaller sample size, we designed a 3-stage machine learning classification architecture, which combines feature selection, dimensionality reduction, and classification into a multistep process. This approach allows for the extraction of the most relevant features while minimizing noise, thereby improving classification performance. Despite the small sample size, our model consistently outperformed conventional single-stage classifiers, highlighting the strength of both the machine learning architecture and the selected human activity features—particularly sleep patterns and activity levels—used for depression classification. This also underscores the adaptability of our model to different data scales, making it a more versatile option for future research where data availability might be limited. While this model shows promising results, however, caution is needed in interpreting these results. Future work should aim to enhance its robustness and generalizability by expanding the dataset. Collecting Wi-Fi–based physical activity data from a larger and more diverse sample would not only improve the model's statistical power but also allow for a more comprehensive evaluation of its performance across different population groups, such as varying age ranges and health conditions. This would be particularly valuable in developing a scalable solution for real-world applications. Additionally, the integration of advanced machine learning techniques, such as deep neural networks or hybrid models combining traditional classifiers with deep learning components, could further enhance classification accuracy.

Our study is distinctive not only in its methodological approach but also in its emphasis on model explainability, a crucial aspect often overlooked in prior research on depression classification. Explainability is essential in health care applications, where understanding the factors driving a model's decision is critical for clinical adoption and trust. By using SHAP and LIME, we were able to dissect the decision-making process of our model and pinpoint the most influential features for classifying depression. Both explainability analyses converge on the identification of sleep interruption features as key predictors in the depression classification of our proposed model. Among these, the "percentage of nights with sleep interruptions," "average sleep duration," and "total number of sleep interruptions" emerged as the primary driving factors. These findings align with existing literature that highlights the strong correlation between sleep disturbances and depression. However, our approach goes a step further by quantifying the impact of these features on the classification outcomes, providing a more nuanced understanding of their role. These findings suggest that future tools for depression assessment may benefit from a stronger focus on sleep quality and patterns and further investigations are required in this regard.

The integration of sensors and AI is transforming health care, yet the application of these technologies in depression classification remains underdeveloped and lacks extensive investigation. This study aimed to create an automated machine learning system for the continuous, remote monitoring and assessment of daily physical activity among older adults in a home setting, with the goal of distinguishing between individuals with and without depression.

In summary, although there were some challenges, our results suggest that using Wi-Fi–based data to capture contextual human activities is a promising and efficient method for classifying depression. The model we developed, leveraging data from Wi-Fi motion sensors, showed strong potential in accurately identifying early signs of depression and paving the way for more advanced and accessible mental health monitoring technologies among community dwelling older adults..

Acknowledgments

This study was funded by the Ministry of Innovation of Quebec, McGill University, Aerial Tech, and Mila—Quebec AI Institute. The authors acknowledge support from these institutions. SAR is Canada Research Chair (Tier II) in Artificial Intelligence and Advanced Digital Primary Health Care, received salary support from a Research Scholar Junior 1 Career Development Award from the Fonds de Recherche du Québec-Santé during part of this project, and her research program is supported by the Natural Sciences Research Council Discovery (grant 2020-05246).

Data Availability

The data analyzed during this study are not publicly available due to restrictions placed by the McGill University Institutional Ethics Committee, but the deidentified data are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization was led by SAR, who established the study's goals, design, and research questions and obtained the funding for the project. The methodology was developed by SAR, and SN. The data collection was done by SAR and VK. Data curation was managed by SAR, SN, and VK. Formal analysis was conducted by SAR and SN. The original draft was written by SAR, and SN. Reviewing and editing were a collaborative effort with all authors. Supervision and overall project leadership were provided by SAR.

Conflicts of Interest

None declared.

References

1. Abd-Alrazaq A, AlSaad R, Shuweihdi F, Ahmed A, Aziz S, Sheikh J. Systematic review and meta-analysis of performance of wearable artificial intelligence in detecting and predicting depression. NPJ Digit Med 2023;6(1):84 [FREE Full text] [doi: 10.1038/s41746-023-00828-5] [Medline: 37147384]
2. Su YA, Si T. Progress and challenges in research of the mechanisms of anhedonia in major depressive disorder. Gen Psychiatr 2022;35(1):e100724 [FREE Full text] [doi: 10.1136/gpsych-2021-100724] [Medline: 35309242]
3. Parra F, Benezeth Y, Yang F. Automatic assessment of emotion dysregulation in American, French, and Tunisian adults and new developments in deep multimodal fusion: cross-sectional study. JMIR Ment Health 2022;9(1):e34333 [FREE Full text] [doi: 10.2196/34333] [Medline: 35072643]
4. Choi J, Lee S, Kim S, Kim D, Kim H. Depressed mood prediction of elderly people with a wearable band. Sensors 2022;22(11):4174 [FREE Full text] [doi: 10.3390/s22114174] [Medline: 35684797]
5. Gao Y, Jia Z, Zhao L, Han S. The effect of activity participation in middle-aged and older people on the trajectory of depression in later life: national cohort study. JMIR Public Health Surveill 2023;9:e44682 [FREE Full text] [doi: 10.2196/44682] [Medline: 36951932]
6. Song S, Seo Y, Hwang S, Kim HY, Kim J. Digital phenotyping of geriatric depression using a community-based digital mental health monitoring platform for socially vulnerable older adults and their community caregivers: 6-week living lab single-arm pilot study. JMIR Mhealth Uhealth 2024;12:e55842 [FREE Full text] [doi: 10.2196/55842] [Medline: 38885033]
7. Colizzi M, Lasalvia A, Ruggeri M. Prevention and early intervention in youth mental health: Is it time for a multidisciplinary and trans-diagnostic model for care? Int J Ment Health Syst 2020;14(1):23 [FREE Full text] [doi: 10.1186/s13033-020-00356-9] [Medline: 32226481]
8. Owen D, Antypas D, Hassoulas A, Pardiñas AF, Espinosa-Anke L, Collados JC. Enabling early health care intervention by detecting depression in users of web-based forums using language models: longitudinal analysis and evaluation. JMIR AI 2023;2:e41205 [FREE Full text] [doi: 10.2196/41205] [Medline: 37525646]
9. Huang S, Wang Y, Li G, Hall BJ, Nyman TJ. Digital mental health interventions for alleviating depression and anxiety during psychotherapy waiting lists: systematic review. JMIR Ment Health 2024;11:e56650 [FREE Full text] [doi: 10.2196/56650] [Medline: 39255015]
10. Smith KM, Renshaw PF, Bilello J. The diagnosis of depression: current and emerging methods. Compr Psychiatry 2013;54(1):1-6 [FREE Full text] [doi: 10.1016/j.comppsych.2012.06.006] [Medline: 22901834]
11. Minaeva O, Riese H, Lamers F, Antypa N, Wichers M, Booij SH. Screening for depression in daily life: development and external validation of a prediction model based on actigraphy and experience sampling method. J Med Internet Res 2020;22(12):e22634 [FREE Full text] [doi: 10.2196/22634] [Medline: 33258783]
12. Glavin D, Grua EM, Nakamura CA, Scazufca M, Ribeiro Dos Santos E, Wong GHY, et al. Patient health questionnaire-9 item pairing predictiveness for prescreening depressive symptomatology: machine learning analysis. JMIR Ment Health 2023;10:e48444 [FREE Full text] [doi: 10.2196/48444] [Medline: 37856186]
13. Camp N, Johnston J, Lewis MGC, Zecca M, Di Nuovo A, Hunter K, et al. Perceptions of in-home monitoring technology for activities of daily living: semistructured interview study with community-dwelling older adults. JMIR Aging 2022;5(2):e33714 [FREE Full text] [doi: 10.2196/33714] [Medline: 35511248]
14. Chan A, Cai J, Qian L, Coutts B, Phan S, Gregson G, et al. In-home positioning for remote home health monitoring in older adults: systematic review. JMIR Aging 2024;7:e57320 [FREE Full text] [doi: 10.2196/57320] [Medline: 39622026]
15. Gianfredi V, Blandi L, Cacitti S, Minelli M, Signorelli C, Amerio A, et al. Depression and objectively measured physical activity: a systematic review and meta-analysis. Int J Environ Res Public Health 2020;17(10):3738 [FREE Full text] [doi: 10.3390/ijerph17103738] [Medline: 32466242]
16. Arrieta H, Rezola-Pardo C, Echeverria I, Iturburu M, Gil SM, Yanguas JJ, et al. Physical activity and fitness are associated with verbal memory, quality of life and depression among nursing home residents: preliminary data of a randomized controlled trial. BMC Geriatr 2018;18(1):80. [doi: 10.1186/s12877-018-0770-y] [Medline: 29580209]
17. Dillon CB, McMahon E, O'Regan G, Perry IJ. Associations between physical behaviour patterns and levels of depressive symptoms, anxiety and well-being in middle-aged adults: a cross-sectional study using isotemporal substitution models. BMJ Open 2018;8(1):e018978 [FREE Full text] [doi: 10.1136/bmjopen-2017-018978] [Medline: 29358436]

18. Jung S, Lee S, Lee S, Bae S, Imaoka M, Harada K, et al. Relationship between physical activity levels and depressive symptoms in community-dwelling older Japanese adults. Geriatr Gerontol Int 2018;18(3):421-427. [doi: 10.1111/ggi.13195] [Medline: 29052928]
19. Ludwig VM, Bayley A, Cook DG, Stahl D, Treasure JL, Asthworth M, et al. Association between depressive symptoms and objectively measured daily step count in individuals at high risk of cardiovascular disease in South London, UK: a cross-sectional study. BMJ Open 2018;8(4):e020942 [FREE Full text] [doi: 10.1136/bmjopen-2017-020942] [Medline: 29654044]
20. Ku PW, Steptoe A, Liao Y, Sun WJ, Chen LJ. Prospective relationship between objectively measured light physical activity and depressive symptoms in later life. Int J Geriatr Psychiatry 2018;33(1):58-65. [doi: 10.1002/gps.4672] [Medline: 28181713]
21. Raudsepp L, Riso EM. Longitudinal association between objectively measured walking and depressive symptoms among Estonian older adults. J Aging Phys Act 2017;25(4):639-645. [doi: 10.1123/japa.2016-0303] [Medline: 28338399]
22. Kumar P, Garg S, Garg A. Assessment of anxiety, depression and stress using machine learning models. Procedia Comput Sci 2020;171:1989-1998. [doi: 10.1016/j.procs.2020.04.213]
23. Adamczyk J, Malawski F. Comparison of manual and automated feature engineering for daily activity classification in mental disorder diagnosis. Comput Inf 2021;40(4):850-879. [doi: 10.31577/cai_2021_4_850]
24. Bai R, Xiao L, Guo Y, Zhu X, Li N, Wang Y, et al. Tracking and monitoring mood stability of patients with major depressive disorder by machine learning models using passive digital data: prospective naturalistic multicenter study. JMIR Mhealth Uhealth 2021;9(3):e24365 [FREE Full text] [doi: 10.2196/24365] [Medline: 33683207]
25. Chikersal P, Doryab A, Tumminia M, Villalba DK, Dutcher JM, Liu X, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing. ACM Trans Comput Hum Interact 2021;28(1):1-41. [doi: 10.1145/3422821]
26. Dai R, Kannampallil T, Zhang J, Lv N, Ma J, Lu C. Multi-task learning for randomized controlled trials. Proc ACM Interact Mob Wearable Ubiquitous Technol 2022;6(2):1-23. [doi: 10.1145/3534591]
27. Griffiths C, da Silva KM, Leathlean C, Jiang H, Ang CS, Searle R. Investigation of physical activity, sleep, and mental health recovery in treatment resistant depression (TRD) patients receiving repetitive transcranial magnetic stimulation (rTMS) treatment. J Affect Disord Rep 2022;8:100337 [FREE Full text] [doi: 10.1016/j.jadr.2022.100337] [Medline: 35619990]
28. Espino-Salinas CH, Galván-Tejada CE, Luna-García H, Gamboa-Rosales H, Celaya-Padilla JM, Zanella-Calzada LA, et al. Two-dimensional convolutional neural network for depression episodes detection in real time using motor activity time series of depresjon dataset. Bioengineering 2022;9(9):458 [FREE Full text] [doi: 10.3390/bioengineering9090458] [Medline: 36135004]
29. Jakobsen P, Garcia-Ceja E, Riegler M, Stabell LA, Nordgreen T, Torresen J, et al. Applying machine learning in motor activity time series of depressed bipolar and unipolar patients compared to healthy controls. PLoS One 2020;15(8):e0231995 [FREE Full text] [doi: 10.1371/journal.pone.0231995] [Medline: 32833958]
30. Jung D, Kim J, Mun KR. Identifying depression in the elderly using gait accelerometry. Annu Int Conf IEEE Eng Med Biol Soc 2022;2022:4946-4949. [doi: 10.1109/EMBC48229.2022.9871877] [Medline: 36086152]
31. Kulam S. Time-series classification with uni-dimensional convolutional neural networks: An experimental comparison with long short-term memory networks [dissertation].: University of Oslo; 2019. URL: https://www.duo.uio.no/handle/10852/73250 [accessed 2024-12-07]
32. Kumar A, Sangwan SR, Arora A, Menon VG. Depress-DCNF: A deep convolutional neuro-fuzzy model for detection of depression episodes using IoMT. Appl Soft Comput 2022;122:108863. [doi: 10.1016/j.asoc.2022.108863]
33. Nguyen DK, Chan CL, Adams Li AH, Phan DV. Deep stacked generalization ensemble learning models in early diagnosis of depression illness from wearable devices data. 2021 Presented at: ICMHI 2021: 2021 5th International Conference on Medical and Health Informatics; May 14-16, 2021; Kyoto, Japan. [doi: 10.1145/3472813.3472815]
34. Opoku Asare K, Moshe I, Terhorst Y, Vega J, Hosio S, Baumeister H, et al. Mood ratings and digital biomarkers from smartphone and wearable data differentiates and predicts depression status: a longitudinal data analysis. Pervasive Mobile Comput 2022;83:101621. [doi: 10.1016/j.pmcj.2022.101621]
35. Park C, Mishra R, Golledge J, Najafi B. Digital biomarkers of physical frailty and frailty phenotypes using sensor-based physical activity and machine learning. Sensors 2021;21(16):5289 [FREE Full text] [doi: 10.3390/s21165289] [Medline: 34450734]
36. Akbari G, Nikkhoo M, Wang L, Chen CPC, Han DS, Lin YH, et al. Frailty level classification of the community elderly using Microsoft Kinect–based skeleton pose: a machine learning approach. Sensors (Basel) 2021;21(12):4017 [FREE Full text] [doi: 10.3390/s21124017] [Medline: 34200838]
37. Utsumi T, Hashimoto M. A development of an early detection system of pre-frailty in senior citizens living inside. 2023 Presented at: 17th International Conference on Ubiquitous Information Management and Communication (IMCOM); Seoul, Korea; January 03-05, 2023. [doi: 10.1109/imcom56909.2023.10035612]

38. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019;21(8):e12785 [FREE Full text] [doi: 10.2196/12785] [Medline: 31471958]
39. Tegou T, Kalamaras I, Tsipouras M, Giannakeas N, Votis K, Tzovaras D. A low-cost indoor activity monitoring system for detecting frailty in older adults. Sensors 2019;19(3):452 [FREE Full text] [doi: 10.3390/s19030452] [Medline: 30678280]
40. Goonawardene N, Tan HP, Tan LB. Unobtrusive detection of frailty in older adults. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population Applications in Health, Assistance, and Entertainment. Cham, Switzerland: Springer International Publishing; 2018.
41. Mohtadifar M, Cheffena M, Pourafzal A. Acoustic- and radio-frequency–based human activity recognition. Sensors 2022;22(9):3125 [FREE Full text] [doi: 10.3390/s22093125] [Medline: 35590815]
42. Wang S, Zhou G. A review on radio based activity recognition. Digit Commun Netw 2015;1(1):20-29. [doi: 10.1016/j.dcan.2015.02.006]
43. Showmik IA, Sanam TF, Imtiaz H. Human activity recognition from Wi-Fi CSI data using principal component-based wavelet CNN. Digit Signal Process 2023;138:104056. [doi: 10.1016/j.dsp.2023.104056]
44. Pulle V, Rozon M, Zandifar A, Yadavalli SH, Asghari VR, Minayi Jalil A, et al. Systems and methods for remote monitoring care using wireless sensing technologies. Justia. URL: https://patents.justia.com/patent/20240027564 [accessed 2024-10-29]
45. Allegue M, Ghourchian N, Rozon M, Carignan E, Hopper M, Gonzalez JF, et al. WiFi motion intelligence: The fundamentals. Aerial Technologies Inc. URL: https://aerial.ai/files/common/The_fundamentals_Aerial_Telefonica_PCP-final_aligned.pdf [accessed 2025-01-14]
46. Vitkovic J, Le C, Lee SL, Clark RA. The contribution of hearing and hearing loss to balance control. Audiol Neurootol 2016;21(4):195-202. [doi: 10.1159/000445100] [Medline: 27251708]
47. Dent E, Kowal P, Hoogendijk EO. Frailty measurement in research and clinical practice: a review. Eur J Intern Med 2016;31:3-10. [doi: 10.1016/j.ejim.2016.03.007] [Medline: 27039014]
48. Rolfson DB, Majumdar SR, Tsuyuki RT, Tahir A, Rockwood K. Validity and reliability of the Edmonton Frail Scale. Age Ageing 2006;35(5):526-529 [FREE Full text] [doi: 10.1093/ageing/afl041] [Medline: 16757522]
49. Perna S, Francis MD, Bologna C, Moncaglieri F, Riva A, Morazzoni P, et al. Performance of Edmonton Frail Scale on frailty assessment: its association with multi-dimensional geriatric conditions assessed with specific screening tools. BMC Geriatr 2017;17(1):2 [FREE Full text] [doi: 10.1186/s12877-016-0382-3] [Medline: 28049443]
50. Yesavage JA. The use of self-rating depression scales in the elderly. In: Poon LW, Crook T, Davis KL, Eisdorfer C, Gurland BJ, Kaszniak AW, et al, editors. Handbook for Clinical Memory Assessment of Older Adults. Washington, DC: American Psychological Association; 1986.
51. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med 2020;26(9):1320-1324 [FREE Full text] [doi: 10.1038/s41591-020-1041-y] [Medline: 32908275]
52. Heo YJ. Population aging and house prices: Who are we calling old? J Econ Ageing 2022;23:100417 [FREE Full text] [doi: 10.1016/j.jeoa.2022.100417]
53. Sarawagi S, Bhamidipaty A. Interactive deduplication using active learning. 2002 Presented at: KDD02: The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; July 23-26, 2002; Edmonton, Canada. [doi: 10.1145/775047.775087]
54. Ghadiri P, Yaffe MJ, Adams AM, Abbasgholizadeh-Rahimi S. Primary care physicians' perceptions of artificial intelligence systems in the care of adolescents' mental health. BMC Prim Care 2024;25(1):215 [FREE Full text] [doi: 10.1186/s12875-024-02417-1] [Medline: 38872128]
55. Rahimi SA, Légaré F, Sharma G, Archambault P, Zomahoun HTV, Chandavong S, et al. Application of artificial intelligence in community-based primary health care: systematic scoping review and critical appraisal. J Med Internet Res 2021;23(9):e29839 [FREE Full text] [doi: 10.2196/29839] [Medline: 34477556]
56. Hall MA. Correlation-Based Feature Selection for Machine Learning [dissertation]. Hamilton, New Zealand: University of Waikato; 1999. URL: https://www.lri.fr/~pierres/donn%E9es/save/these/articles/lpr-queue/hall99correlationbased.pdf [accessed 2025-01-14]
57. Abdi H, Williams LJ. Principal component analysis. WIREs Computational Stats 2010;2(4):433-459. [doi: 10.1002/wics.101]
58. Rish I. An empirical study of the naive Bayes classifier. 2001 Presented at: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence; August 4, 2001; Seattle, WA URL: https://www.semanticscholar.org/paper/An-empirical-study-of-the-naive-Bayes-classifier-Watson/2825733f97124013e8841b3f8a0f5bd4ee4af88a
59. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA URL: https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
60. Marcano-Cedeño A, Quintanilla-Domínguez J, Cortina-Januchs MG, Andina D. Feature selection using sequential forward selection and classification applying artificial metaplasticity neural network. 2010 Presented at: IECON 2010—36th Annual Conference on IEEE Industrial Electronics Society; November 7-10, 2010; Glendale, AZ. [doi: 10.1109/iecon.2010.5675075]

61. Fodor IK. A survey of dimension reduction techniques. Office of Scientific and Technical Information. 2002. URL: https://www.osti.gov/biblio/15002155 [accessed 2022-05-09]
62. Jr DWH, Lemeshow S. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; 2004.
63. Peltola T. Local interpretable model-agnostic explanations of Bayesian predictive models via Kullback-Leibler projections. ArXiv. Preprint posted online on October 5, 2018 2018. [doi: 10.48550/arXiv.1810.02678]
64. Vergara JR, Estévez PA. A review of feature selection methods based on mutual information. Neural Comput Appl 2013;24(1):175-186. [doi: 10.1007/s00521-013-1368-0]
65. Balakrishnama S, Ganapathiraju A. Linear Discriminant Analysis—A Brief Tutorial [dissertation].: Mississippi State University; 1995. URL: https://www.semanticscholar.org/paper/LINEAR-DISCRIMINANT-ANALYSIS-A-BRIEF-TUTORIAL-Balakrishnama-Ganapathiraju/163bb761cd7c5ca8c786bf4816a10308f1396ffb [accessed 2024-10-29]
66. Peterson LE. K-nearest neighbor. Scholarpedia 2009;10(2):1883. [doi: 10.4249/scholarpedia.1883]
67. Tariq MA. A study on comparative analysis of feature selection algorithms for students grades prediction. J Inf Organ Sci 2024;48(1):133-147. [doi: 10.31341/jios.48.1.7]
68. Cao LJ, Chua KS, Chong WK, Lee HP, Gu QM. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing 2003;55(1-2):321-336. [doi: 10.1016/s0925-2312(03)00433-8]
69. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Appl 1998;13(4):18-28. [doi: 10.1016/j.aca.2011.07.027] [Medline: 21889629]
70. Granitto PM, Furlanello C, Biasioli F, Gasperi F. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. Chemometrics Intell Lab Syst 2006;83(2):83-90. [doi: 10.1016/j.chemolab.2006.01.007]
71. Song Y, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry 2015;27(2):130-135 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.215044] [Medline: 26120265]
72. Breiman L. Random forests. Machine Learning 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
73. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann. Statist 2001;29(5):1189-1232. [doi: 10.1214/aos/1013203451]
74. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. xgboost: Extreme Gradient Boosting. R Foundation for Statistical Computing. 2014. URL: https://cran.r-project.org/web/packages/xgboost/index.html [accessed 2024-07-24]
75. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A highly efficient gradient boosting decision tree. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA.
76. Ruta D, Gabrys B. Classifier selection for majority voting. Inf Fusion 2005;6(1):63-81. [doi: 10.1016/j.inffus.2004.04.008]
77. Zareapoor M, Shamsolmoali P. Application of credit card fraud detection: based on bagging ensemble classifier. Procedia Comput Sci 2015;48:679-685. [doi: 10.1016/j.procs.2015.04.201]
78. Hastie T, Rosset S, Zhu J, Zou H. Multi-class AdaBoost. Stat Interface 2009;2(3):349-360. [doi: 10.4310/sii.2009.v2.n3.a8]
79. Khalid S, Khalil T, Nasreen S. A survey of feature selection and feature extraction techniques in machine learning. Sci Inf Conf 2014:024-032 [FREE Full text] [doi: 10.14308/ite000494]
80. Tsai CF, Sung YT. Ensemble feature selection in high dimension, low sample size datasets: parallel and serial combination approaches. Knowl Based Syst 2020;203:106097. [doi: 10.1016/j.knosys.2020.106097]
81. Soares I, Dias J, Rocha H, Do Carmo Lopes M, Ferreira B. Feature selection in small databases: a medical-case study. 2016 Presented at: XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016; March 31-April 2, 2016; Paphos, Cyprus. [doi: 10.1007/978-3-319-32703-7_158]
82. Lever J, Krzywinski M, Altman N. Principal component analysis. Nature Methods 2017;14(7):641-642. [doi: 10.1038/nmeth.4346]
83. Saed AAA, Jaharadak AA. Implementation with performance evaluation of decision tree classifier for uncertain data. Lit Rev 2022;5(5):125-132. [doi: 10.1017/cbo9780511810114.020]
84. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res 2012;13(1):281-305 [FREE Full text]
85. Frazier P. A tutorial on Bayesian optimization. ArXiv. Preprint posted online July 8, 2018 2018 [FREE Full text] [doi: 10.1287/educ.2018.0188]
86. Li L, Jamieson KG, DeSalvo G, Rostamizadeh A, Talwalkar A. Hyperband: A novel bandit-based approach to hyperparameter optimization. J Mach Learn Res 2016:6765-6816 [FREE Full text]
87. Shreffler J, Huecker MR. Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios. Treasure Island, FL: StatPearls Publishing; 2024.
88. Liu W, Jiang Y, Zhang J, Ma Z. RFARN: Retinal vessel segmentation based on reverse fusion attention residual network. PLoS One 2021;16(12):e0257256 [FREE Full text] [doi: 10.1371/journal.pone.0257256] [Medline: 34860847]
89. Goodyear MDE, Krleza-Jeric K, Lemmens T. The Declaration of Helsinki. BMJ 2007;335(7621):624-625 [FREE Full text] [doi: 10.1136/bmj.39339.610000.BE] [Medline: 17901471]
90. Tri-Council policy statement: ethical conduct for research involving humans—TCPS 2 (2022). Government of Canada. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html [accessed 2024-10-29]

91. Shayan Nejadshamsi. AID project - Depression Github Repository. URL: https://github.com/rahimi-s-lab/HOPE-model [accessed 2025-02-07]
92. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2015;12(85):2825-2830. [doi: 10.1145/2786984.2786995]
93. Akyol Y, Durmus D, Dogan C, Bek Y, Canturk F. Quality of life and level of depressive symptoms in the geriatric population. Turk J Rheumatol 2010;25(4):165-173. [doi: 10.5152/tjr.2010.23]
94. de Jonge P, Ormel J, Slaets JPJ, Kempen GIJM, Ranchor AV, van Jaarsveld CHM, et al. Depressive symptoms in elderly patients predict poor adjustment after somatic events. Am J Geriatr Psychiatry 2004;12(1):57-64. [Medline: 14729560]
95. Helvik AS, Engedal K, Krokstad S, Stordal E, Selbæk G. A comparison of depressive symptoms in elderly medical inpatients and the elderly in a population-based health study (the Nord-Trøndelag Health Study 3). Nord J Psychiatry 2012;66(3):189-197. [doi: 10.3109/08039488.2011.614958] [Medline: 21958367]
96. Ordu Gokkaya NK, Gokce-Kutsal Y, Borman P, Ceceli E, Dogan A, Eyigor S, et al. Pain and quality of life (QoL) in elderly: the Turkish experience. Arch Gerontol Geriatr 2012;55(2):357-362. [doi: 10.1016/j.archger.2011.10.019] [Medline: 22104759]
97. Hsu HC. Physical function trajectories, depressive symptoms, and life satisfaction among the elderly in Taiwan. Aging Ment Health 2009;13(2):202-212. [doi: 10.1080/13607860802342201] [Medline: 19347687]
98. Lue BH, Chen LJ, Wu SC. Health, financial stresses, and life satisfaction affecting late-life depression among older adults: a nationwide, longitudinal survey in taiwan. Arch Gerontol Geriatr 2010;50:S34-S38. [doi: 10.1016/S0167-4943(10)70010-8] [Medline: 20171454]
99. Feng L, Yap KB, Ng TP. Depressive symptoms in older adults with chronic kidney disease: mortality, quality of life outcomes, and correlates. Am J Geriatr Psychiatry 2013;21(6):570-579. [doi: 10.1016/j.jagp.2012.12.020] [Medline: 23567405]
100. Van der Weele GM, Gussekloo J, De Waal MWM, De Craen AJM, Van der Mast RC. Co-occurrence of depression and anxiety in elderly subjects aged 90 years and its relationship with functional status, quality of life and mortality. Int J Geriatr Psychiatry 2009;24(6):595-601. [doi: 10.1002/gps.2162] [Medline: 19031476]
101. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]
102. Byrne JEM, Bullock B, Brydon A, Murray G. A psychometric investigation of the sleep, circadian rhythms, and mood (SCRAM) questionnaire. Chronobiol Int 2019;36(2):265-275. [doi: 10.1080/07420528.2018.1533850] [Medline: 30395721]
103. Difrancesco S, Lamers F, Riese H, Merikangas KR, Beekman ATF, van Hemert AM, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: a 2-week ambulatory assessment study. Depress Anxiety 2019;36(10):975-986 [FREE Full text] [doi: 10.1002/da.22949] [Medline: 31348850]
104. Ghandeharioun A, Fedor S, Sangermano L, Ionescu D, Alpert J, Dale C, et al. Objective assessment of depressive symptoms with machine learning and wearable sensors data. 2017 Presented at: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII); October 23-26, 2017; San Antonio, TX. [doi: 10.1109/acii.2017.8273620]
105. Jacobson NC, Weingarden H, Wilhelm S. Using digital phenotyping to accurately detect depression severity. J Nerv Ment Dis 2019;207(10):893-896. [doi: 10.1097/NMD.0000000000001042] [Medline: 31596769]
106. Vallance JK, Eurich D, Lavallee C, Johnson ST. Daily pedometer steps among older men: associations with health-related quality of life and psychosocial health. Am J Health Promot 2013;27(5):294-298. [doi: 10.4278/ajhp.120316-QUAN-145] [Medline: 23402230]
107. Aprahamian I, Borges MK, Hanssen DJC, Jeuring HW, Oude Voshaar RC. The frail depressed patient: a narrative review on treatment challenges. Clin Interv Aging 2022;17:979-990 [FREE Full text] [doi: 10.2147/CIA.S328432] [Medline: 35770239]
108. Lohman MC, Mezuk B. Frailty and depression: comorbidity in the context of imperfect measurement. J Am Geriatr Soc 2013;61(3):474 [FREE Full text] [doi: 10.1111/jgs.12138] [Medline: 23496193]
109. Vaughan L, Corbin AL, Goveas JS. Depression and frailty in later life: a systematic review. Clin Interv Aging 2015;10:1947-1958 [FREE Full text] [doi: 10.2147/CIA.S69632] [Medline: 26719681]

Abbreviations

AI: artificial intelligence
CNN: convolutional neural network
CSI: channel state information
DT: decision tree
EFS: Edmonton Frailty Scale
GDS: Geriatric Depression Scale
LIME: local interpretable model-agnostic explanations
LR: logistic regression

MI-CLAIM: Minimum Information About Clinical Artificial Intelligence Modeling
PCA: principal component analysis
RF: random forest
RSSI: received signal strength indicator
SFS: sequential forward selection
SHAP: Shapely addictive explanations
SVM: support vector machine

Edited by Y Du; submitted 18.10.24; peer-reviewed by PH Liao, C Silva; comments to author 05.12.24; revised version received 12.12.24; accepted 19.12.24; published 03.03.25.

Please cite as:
Nejadshamsi S, Karami V, Ghourchian N, Armanfard N, Bergman H, Grad R, Wilchesky M, Khanassov V, Vedel I, Abbasgholizadeh Rahimi S
Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study
JMIR Aging 2025;8:e67715
URL: https://aging.jmir.org/2025/1/e67715
doi:10.2196/67715
PMID:40053734

Original Paper

Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study

Ming De Lim[1], BCS; Tee Connie[1], PhD; Michael Kah Ong Goh[1], PhD; Nor 'Izzati Saedon[2], MBBS

[1]Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia

[2]Department of Medicine, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:
Tee Connie, PhD
Faculty of Information Science and Technology
Multimedia University
Jalan Ayer Keroh Lama
Melaka, 75450
Malaysia
Phone: 60 62523592
Fax: 60 62318840
Email: tee.connie@mmu.edu.my

Abstract

Background: Parkinson disease (PD) is a progressive neurodegenerative disorder that affects motor coordination, leading to gait abnormalities. Early detection of PD is crucial for effective management and treatment. Traditional diagnostic methods often require invasive procedures or are performed when the disease has significantly progressed. Therefore, there is a need for noninvasive techniques that can identify early motor symptoms, particularly those related to gait.

Objective: The study aimed to develop a noninvasive approach for the early detection of PD by analyzing model-based gait features. The primary focus is on identifying subtle gait abnormalities associated with PD using kinematic characteristics.

Methods: Data were collected through controlled video recordings of participants performing the timed up and go (TUG) assessment, with particular emphasis on the turning phase. The kinematic features analyzed include shoulder distance, step length, stride length, knee and hip angles, leg and arm symmetry, and trunk angles. These features were processed using advanced filtering techniques and analyzed through machine learning methods to distinguish between normal and PD-affected gait patterns.

Results: The analysis of kinematic features during the turning phase of the TUG assessment revealed that individuals with PD exhibited subtle gait abnormalities, such as freezing of gait, reduced step length, and asymmetrical movements. The model-based features proved effective in differentiating between normal and PD-affected gait, demonstrating the potential of this approach in early detection.

Conclusions: This study presents a promising noninvasive method for the early detection of PD by analyzing specific gait features during the turning phase of the TUG assessment. The findings suggest that this approach could serve as a sensitive and accurate tool for diagnosing and monitoring PD, potentially leading to earlier intervention and improved patient outcomes.

(JMIR Aging 2025;8:e65629) doi:10.2196/65629

KEYWORDS

model-based features; gait analysis; Parkinson disease; computer vision; support vector machine

Introduction

Background

Parkinson disease (PD) is a common neurological disease that affects millions of individuals worldwide. This disorder gradually impairs a person's ability to move their body, resulting in a variety of crippling symptoms, including tremors, stiff muscles, and sluggish motions. To effectively manage PD and provide timely measures, accurate and early diagnosis is essential.

Traditionally, clinical evaluations conducted by medical professionals, which can be arbitrary and inconsistent, have

been used to diagnose PD. In distant or underdeveloped areas, access to specialized PD care may be restricted, delaying diagnosis and therapy. The prevalence of PD is predicted to rise as the world's population ages, placing additional demand on health care resources and highlighting the need for easily available and effective diagnostic methods.

In this regard, current technology has shown promise in addressing the difficulties associated with PD diagnosis, especially in the areas of machine learning and deep learning. The limits of conventional clinical procedures can be overcome with the promise of early and objective identification provided by machine learning techniques.

This study investigated the use of machine learning for PD identification with a focus on gait characteristics. The goal was to develop a simple and noninvasive method for screening PD symptoms early through analyzing gait patterns.

In this study, kinematic features during the turning phase of the timed up and go (TUG) assessment were extracted and analyzed. Gait variabilities occurring during body turning are more easily identified, as turning involves complex motor coordination and balance adjustments, making it a challenging movement. This increased complexity can accentuate subtle abnormalities in gait patterns that might not be as apparent during straight walking. Turning is particularly difficult for individuals with PD, as it requires precise control and stability, often revealing difficulties such as freezing of gait (FoG), reduced step length, and asymmetrical movements. Therefore, analyzing the turning phase allows for a more sensitive and accurate detection of PD symptoms, providing a reliable indicator of whether a participant exhibits PD-related gait abnormalities. This focus enhances the effectiveness of the diagnostic tool, which contributes to the facilitation of early and accurate identification of PD.

Conventional Methods

Challa et al [1] proposed an advanced predictive model for PD by using machine learning algorithms, such as multilayer perceptron, boosted logistic regression, random forest, and BayesNet. Their investigation applied the Parkinson Progression Markers Initiative dataset, an extensive dataset containing data from patients with PD and healthy participants. The experimental results showcased remarkable performance improvements over existing methods, boasting accuracy rates of 96.09% for training and 95.45% for testing in the multilayer perceptron algorithm, 96.5854% for training and 96.02% for testing in the BayesNet algorithm, 95.45% for training and 94.87% for testing in the random forest algorithm, and the highest accuracy achieved by the boosted logistic regression algorithm with 97.159% for training and 96.97% for testing. The area under the curve (AUC) of the receiver operating characteristic curve reached an impressive 98.9% for the boosted logistic regression algorithm, emphasizing its robust predictive capabilities. This research represents a significant stride in health care, providing a reliable model for early PD prediction, crucial for timely diagnosis and intervention in addressing this global public health challenge.

In 2019, Polat [2] investigated the recognition of FoG cases in individuals with PD, using a logistic regression classifier trained and tested on a dataset comprising 16 samples. The study meticulously assessed the classifier's performance using a comprehensive set of 10 performance measures, such as accuracy, miss rate, false discovery rate, false positive rate, false omission rate, sensitivity, specificity, precision, and negative predictive value. Impressively, the logistic regression classifier exhibited a noteworthy accuracy of 81.3% in accurately classifying FoG cases. The research further compared the performance of linear regression with 4 alternative models: linear support vector machine (SVM), quadratic SVM, cubic SVM, and k-nearest neighbors (KNN), revealing that the proposed logistic regression model surpassed its counterparts with the highest accuracy of 81.3% in classifying FoG datasets for individuals with PD. This outcome underscores the superior performance of the logistic regression model in the context of FoG classification.

Vidya and Sasikumar [3] conducted a comprehensive study on the application of multiclass SVM in using gait analysis to identify and grade the severity of PD. The researchers used a publicly accessible dataset containing Vertical Ground Reaction Force Sensors and implemented kinematic analysis to extract spatiotemporal features crucial for the diagnostic process. Their suggested framework included a multiregression strategy to normalize gait time series data and a correlation-based feature selection method. A total of 4 distinct SVM kernel functions such as linear, Gaussian, quadratic, and cubic were rigorously evaluated across 3 different walking tests to gauge their performance. Impressively, the quadratic SVM classifier emerged as the most effective, achieving an outstanding average accuracy of 98.65%. This result surpassed existing state-of-the-art methods, showcasing the robustness and efficacy of the proposed SVM-based approach for PD diagnosis and severity rating.

Moreover, Fang [4] performed a study focusing on predicting PD through the application of machine learning techniques. The research extensively compared the accuracy and recall of 3 distinct algorithms: KNN, random forest, and naive Bayesian. Addressing the inherent limitations of equal weighting in traditional KNN, the study introduced an entropy weight method to enhance KNN's performance, specifically mitigating equal-weighting issues. In addition, the research delved into a voice-based Unified Parkinson Disease Rating Scale (UPDRS) prediction scheme, leveraging algorithms to effectively predict UPDRS scores from voice data of patients with PD. Notably, the study used the University of California Irvine dataset, showcasing that the refined KNN algorithm surpassed its traditional counterpart, achieving a notable accuracy rate increase from 91.8% to 93.8%. This research contributes significantly to the advancement of PD prediction methodologies, particularly emphasizing the pivotal role of improved weighting mechanisms in enhancing algorithmic accuracy.

In addition, Gundala et al [5] conducted a comprehensive study using the random forest algorithm for the recognition of PD using the spiral handwritten dataset. The machine learning technique, widely recognized for its efficacy in processing handwritten designs, was applied by partitioning the dataset into subgroups on the basis of features and constructing decision

trees for each feature. The dataset, sourced from the Kaggle website, provided the foundation for training the random forest model. Notably, the algorithm's strength lies in its ability to combine outputs from multiple decision trees through majority voting, resulting in a remarkable accuracy rate of 91%. The study emphasizes the robustness of the random forest algorithm in accurately identifying PD on the basis of the features extracted from handwritten drawings. The results underscore the algorithm's effectiveness in leveraging the diversity of decision trees for enhanced predictive accuracy in the context of PD diagnosis.

Moreover, Govindu and Palwe [6] analyzed Multidimensional Voice Program audio data collected from both patients with PD and healthy individuals. Among the machine learning models evaluated, the random forest classifier emerged as the most effective, achieving a detection accuracy of 91.83% and a sensitivity of 0.95. This model outperformed other techniques, such as SVM, KNN and logistic regression, which were also assessed for their classification capabilities. The superior performance of the random forest classifier highlights its robustness and reliability for detecting PD in its early stages, demonstrating the promise of machine learning in enhancing diagnostic accuracy.

Deep Learning Methods

Pereira et al [7] proposed an innovative method for early identification of PD using a convolutional neural network (CNN) trained on handwritten dynamics data obtained by a smart pen. The CNN effectively learned relevant features from the signals generated during individual exams, enabling discrimination between individuals with and without PD on the basis of these learned features. Experimental results showcased the superiority of the CNN over raw data, with the ImageNet architecture, using 128 times 128 images and a 75% training dataset split, yielding the best overall accuracy of 83.77%. Despite these promising results, the study acknowledged challenges in achieving consistent recognition rates over control individuals, emphasizing the need for further refinement in the proposed approach.

Later, Grover et al [8] introduced a deep learning methodology using a deep neural network (DNN) constructed with TensorFlow and Keras, containing 3 hidden layers with 10, 20, and 10 neurons each and an input layer with 16 units, concluding with an output layer representing the classes *severe* and *nonsevere*. The DNN was trained on the Parkinson telemonitoring voice dataset from the University of California Irvine machine learning repository, comprising biomedical voice measurements from 42 participants. The experiments focused on predicting the severity of PD based on total UPDRS and motor UPDRS scores. The DNN exhibited substantial accuracy improvements over previous research, achieving a classification accuracy of 94.4422% for total UPDRS and 83.367% for motor UPDRS on the training dataset. While test dataset results were comparatively lower, with 62.7335% accuracy for total UPDRS and 81.6667% for motor UPDRS, the proposed DNN classifier showcased enhanced performance compared to previous studies, emphasizing its potential for accurate severity prediction in PD.

Aal et al [9] conducted an optimized approach for early PD detection, using speech features extracted from 2 datasets, dataset 1 and dataset 2, containing recordings from both healthy individuals and patients with PD. Using Mel-frequency cepstral coefficients and delta Mel-frequency cepstral coefficients, the authors proposed a deep learning model combining a recurrent neural network (RNN) with a long short-term memory (LSTM) layer. Comparative evaluation of alternative ML techniques, including SVM, KNN and RNN with stochastic gradient descent, revealed superior performance of the proposed RNN-LSTM model, optimized with the adaptive moment estimation optimizer. The model exhibited remarkable testing accuracy rates of 95.8% on dataset 1 and 90.24% on DS2, accompanied by high recall, precision, and F_1-score on both datasets. Aal et al [9] further demonstrated the model's superiority over existing methods for PD detection, solidifying its potential as an effective tool for early diagnosis.

Apart from that, Ouhmida et al [10] delved into the early recognition of voice-based PD through the application of CNN and artificial neural network (ANN). Using 2 distinct datasets, the study showcased the superior accuracy of CNN over ANN in their experiments. Dataset 1 encompassed 195 voice recordings from 31 individuals, while dataset 2 featured 240 recordings from 80 participants, with a balanced distribution of 40 patients with PD and 40 healthy individuals. The deep learning models, specifically CNN and ANN, underwent training and testing on both datasets, revealing CNN's accuracy rates of 93.10% and 88.89% for datasets 1 and 2, respectively. In contrast, the ANN model achieved slightly lower accuracies, with 82.76% for dataset 1 and 72.22% for dataset 2. Ouhmida et al [10] also highlighted the intricate layers comprising each model, with the ANN model featuring 2 hidden layers and the CNN model incorporating convolution, normalization, activation, softmax, and classification layers. The paper expressed an intent to extend the research by exploring additional deep learning methods and implementing a hybrid system integrating diverse techniques and datasets.

Biswas et al [11] proposed an approach for the early recognition of PD by proposing 2 distinct deep learning models tailored for hand-drawn graphics. The first model, a 2D CNN, processed preprocessed images of spirals, circles, and meanders as input, achieving notable accuracies of 83.6% on circles, 61.5% on spirals, and 67.8% on meanders. Another model, an innovative LSTM model, operated on timeline-series signals and demonstrated an overall accuracy of 0.78. Leveraging the NewHandPD dataset for training and testing, the authors conducted 4 comprehensive experiments, presenting the results. The study posited that early PD detection through these advanced models could potentially enhance treatment outcomes and elevate the overall quality of life for affected individuals.

Apart from that, Khaskhoussy and Ayed [12] evaluate the effectiveness of SVM CNN for classifying data obtained from speech tasks. Total 2 types of input data were analyzed: the raw speech signal values and i-vector features of dimensions 100, 200, and 300. The classification performance was assessed using 5 evaluation metrics: accuracy, precision, recall or sensitivity, specificity, and F_1-score. For a test dataset of 28 participants,

the approach achieved outstanding results, including 100% accuracy, a precision of 0.99, a recall of 0.98, a specificity of 0.96, and an F_1-score of 0.98. A table summarizing the state-of-the-art methods is presented in Table 1.

Table 1. A summary of the state-of-the-art methods.

Study	Methods	Input type	Dataset	Performance
Challa et al [1]	MLP[a], BayesNet, RF[b], and boosted logistic regression	Nonmotor symptoms	PPMI[c] dataset	Best overall accuracy: 96.97%
Polat [2]	Logistic regression	FoG[d]	FoG dataset (Parkinson disease)	81.3%
Vidya and Sasikumar [3]	SVM[e]	Gait features	Gait analysis dataset	98.65%
Fang [4]	KNN[f]	Voice records	UCI[g] dataset	93.8%
Gundala et al [5]	RF	Handwritten drawings	Kaggle handwritten drawings dataset	91%
Pereira et al [7]	CNN[h]	Handwritten dynamics	Public dataset of handwritten dynamics extracted by a smart pen	Best overall accuracy: 83.77%
Grover et al [8]	DNN[i]	Voice dataset	Parkinson telemonitoring voice dataset	Total UPDRS[j]: train 94.44%, test 62.73% and motor UPDRS: train 83.37%, test 81.67%
Aal et al [9]	RNN[k]-LSTM[l]	Speech features	Dataset 1 and dataset 1[2]	Dataset 1: 95.8% and dataset 12: 90.24%
Ouhmida et al [10]	CNN ANN[m]	Voice	Dataset 1, dataset 2	Dataset 1: 93.10% and dataset 2: 88.89%
Biswas et al [11]	LSTM	Hand drawings	NewHandPD	78.7%

[a]MLP: multilayer perceptron.
[b]RF: random forest.
[c]PPMI: Parkinson Progression Markers Initiative.
[d]FoG: freezing of gait.
[e]SVM: support vector machine.
[f]KNN: k-nearest neighbors.
[g]UCI: University of California Irvine
[h]CNN: convolutional neural network.
[i]DNN: deep neural network.
[j]UPDRS: Unified Parkinson Disease Rating Scale.
[k]RNN-: recurrent neural network.
[l]LSTM: long short-term memory.
[m]ANN: artificial neural network.

Methods

Overview

The study used a methodical approach that started with data collection from young adults, older adults, and patients with PD, followed by obtaining consent, particularly from those diagnosed with the disease. The TUG assessment was a key part of the data collection process. Video enhancements and preprocessing were performed to enhance the quality of the videos. After that, the key points on the human body were obtained using a human pose estimation technique. Next, features such as shoulder distance, step and stride lengths, cadence, and speed were extracted to analyze the gait patterns. The Butterworth filter was applied to refine the data, and peaks were identified to calculate steps and turning durations. Finally, SVM [2] was used to distinguish between the different groups based on the extracted features, aiming to improve the prediction and analysis of PD symptoms.

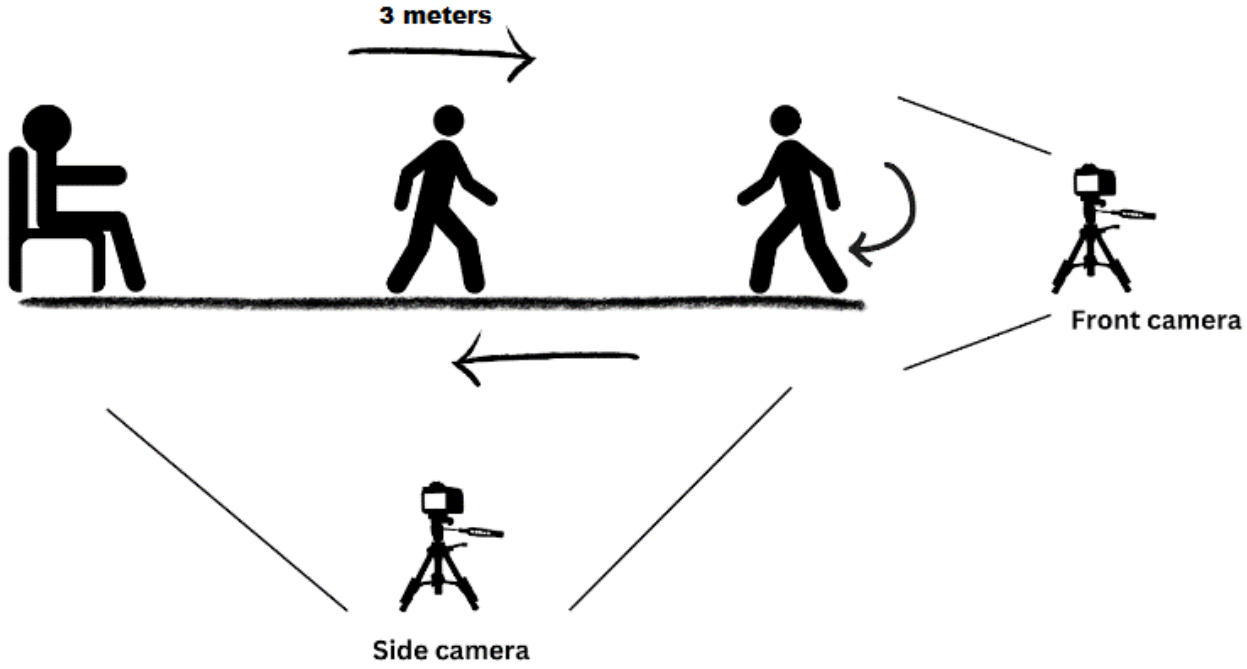
Dataset

The self-collected dataset comprised video recordings of 28 individuals performing the TUG assessment, a standard test used to evaluate mobility and balance. The dataset comprised 3 distinct cohorts: young adults aged between 21 and 33 years, older individuals aged >60 years, and older individuals diagnosed with PD. Consents were obtained from the participants before data collection. The videos captured various gait patterns, providing a comprehensive set of data for analyzing kinematic features, such as step length, stride length, and joint angles. This dataset allowed for controlled conditions and detailed annotations, ensuring high-quality data for feature extraction and analysis.

TUG Assessment

Figure 1 shows a TUG workflow, outlining the sequential steps involved in the activity. The process began with the individual seated at the designated starting point. Upon initiating the recording, the participant stood up and proceeded to walk a distance of 3 meters in a typical manner. Upon reaching the 3-meter mark, the individual executed a turn, walking back to the starting point, where they concluded the TUG process by returning to a seated position. This comprehensive description encapsulated the entire TUG procedure, providing a clear understanding of the task's progression. To enhance data capture, 2 phones were used, one on the front and the other on the side.

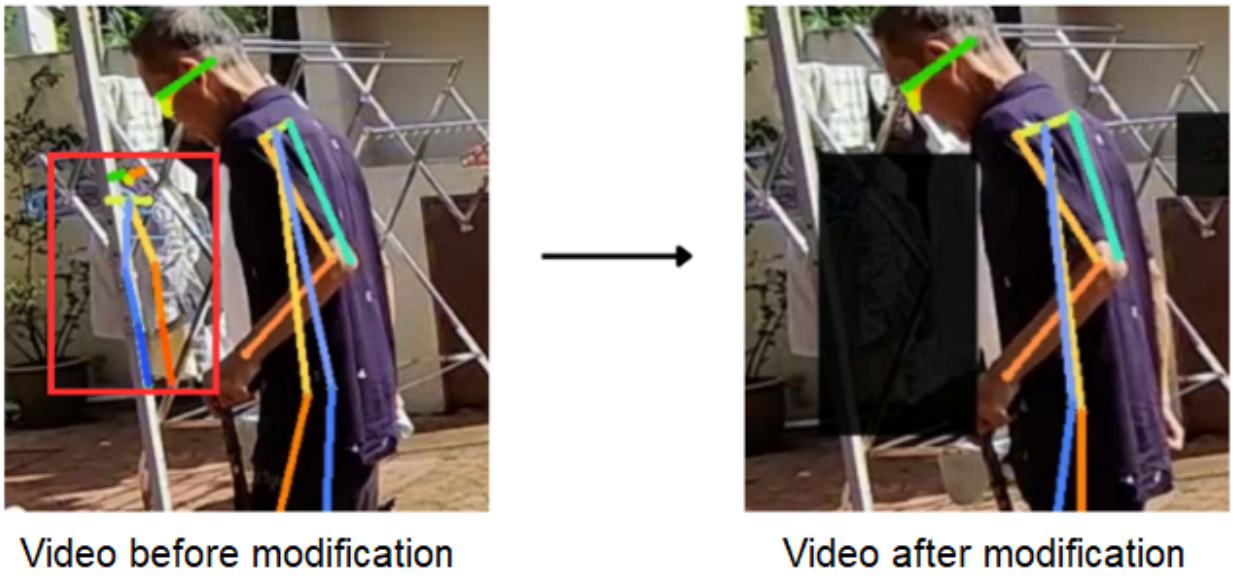
Figure 1. Timed up and go (TUG) assessment shooting site configuration plan.



Video Enhancement

The presence of background clutter in the original video introduces noise, which significantly affects the accuracy of letter pose extraction. Applying human pose estimation technique directly to unprocessed video may result in inaccurate keypoint coordinates, affecting the reliability of analysis results. To alleviate this problem, a preprocessing step was crucial. Specifically, the video was subjected to a customized processing method in which a black block was used to mask certain areas susceptible to noise. This strategic approach ensured that no noise interferes with the subsequent process, resulting in more precise and reliable coordination extraction. Figure 2 shows an example of enhancement performed on the video images.

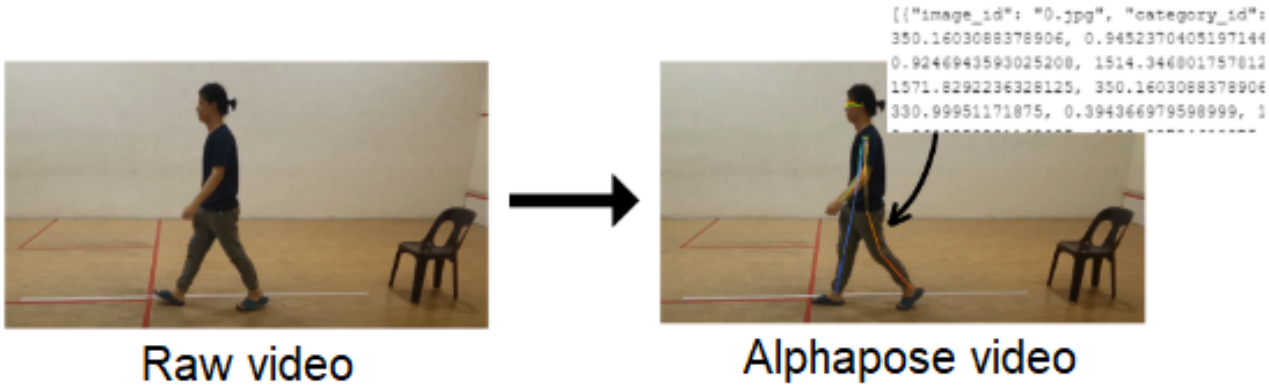
Figure 2. Video enhancement.



Human Pose Estimation

In this study, AlphaPose was used to identify 17 key points from the human body, each corresponding to a specific anatomical location. The process yielded a JSON file containing coefficients for 17 key points. It is important to note that the JSON file, at this stage, has not undergone data processing. The raw file was populated with various elements, including image_id, category_id, key points, and scores within each set. The image_id represented the frame of the video, while the category_id served to identify the object (set as 1 for a person). The key points section contains coordinates for body part locations and corresponding detection confidence, formatted as x1, y1, c1, x2, y2, c2, and so forth, with "c" denoting the confidence score (Figure 3).

Figure 3. Gait analysis using AlphaPose.



Data Extraction

In the data preprocessing phase, the first step involved loading the JSON file and carefully filtering key points and image IDs. In addition, confidence scores were eliminated. The next focus was to mark the 17 identified body key points, including nose, left eye (LEye), right eye (REye), left ear (LEar), right ear (REar), left shoulder (LShoulder), right shoulder (RShoulder), left elbow (LElbow), right elbow (RElbow), left wrist (LWrist), right wrist (RWrist), left hip (LHip), right hip (RHip), left knee (LKnee), right knee (Rknee), left ankle (LAnkle), and right ankle (RAnkle). Next, according to the specific video file name, key point coordinates for each image ID (frame) were marked, ranging from Nose_x and Nose_y to RAnkle_x and RAnkle_y.

The results from the data preprocessing stage boasted a comprehensive structure, encompassing 36 columns that capture key information. Each column was curated to provide a detailed representation of the dataset. Starting with the file_name, which specified the associated video file, and the image_id indicating the frame or image ID corresponding to each set of key points, the subsequent 34 columns focus on the x and y coordinates of 17 distinct body key points.

Frame Segmentation

Frame segmentation is a process to assign a time interval to each frame in the video which helps to determine the occurrence of a turning event. The total duration of the video is divided by the total number of frames (rows). This calculation yields the time interval per frame, assuming a constant frame rate throughout the video. With this interval calculated, a new column named "time_duration" is added to the DataFrame. Each row in the "time_duration" column is populated with the cumulative time, calculated as the frame's index in the DataFrame multiplied by the time interval. This method provided a time stamp for each frame, which is essential for synchronizing the data with the video and analyzing the timing of the detected body turning events. The pseudocode presented in Figure 4 describes the step-by-step procedure for frame segmentation.

Figure 4. Pseudocode describing the step-by-step procedure for frame segmentation.

Algorithm 1: Pseudocode for Frame Segmentation

Data: *df*
Result: *turning_events*

```
function frame_segmentation(DataFrame df)
  for each row in df:
    shoulder_distance = sqrt((RShoulder_x − LShoulder_x)^2 + (RShoulder_y − LShoulder_y)^2)
    set df[shoulder_distance] to shoulder_distance
  end for

  for each row in df:
    if df[shoulder_distance] > threshold then
      set df[turning_event] to TRUE
    else
    set df[turning_event] to FALSE
  end if
 end for

 Return turning_events
End function

# Main Process
DataFrame df = Load data into a DataFrame
Float distance_threshold = Set the distance threshold

DataFrame turning_events = detect_turning_events(df)

Integer total_rows = Length of DataFrame df  # Total number of rows
Float total_duration_seconds = Set the total duration of the video in seconds

Float time_interval = total_duration_seconds / total_rows

For each row in DataFrame turning_events:
   Calculate time_duration = index_of_row * time_interval
   Add time_duration to DataFrame turning_events
```

Noise Filtering

The Butterworth filter was used as the noise filtering tool in this study. The Butterworth filter is a type of signal processing filter that plays a pivotal role in enhancing the clarity of movement coordinate data extracted from videos. By applying the Butterworth filter to the raw coordinates, high-frequency noise is effectively attenuated, resulting in a smoother trajectory on the graph. This smoothed representation provides a clearer visualization of the participant's movements, reducing interference from irrelevant fluctuations.

Combined with the Butterworth filter, a peak detection algorithm was used to identify important points on the graph. These peaks correspond to key events in motion, such as steps taken by the participant. The peak identification process helped in extracting basic features such as the number of steps throughout the video. This kind of feature extraction is particularly valuable for specific PD aspects of analysis and quantification. The Butterworth filter is expressed using equation 1:

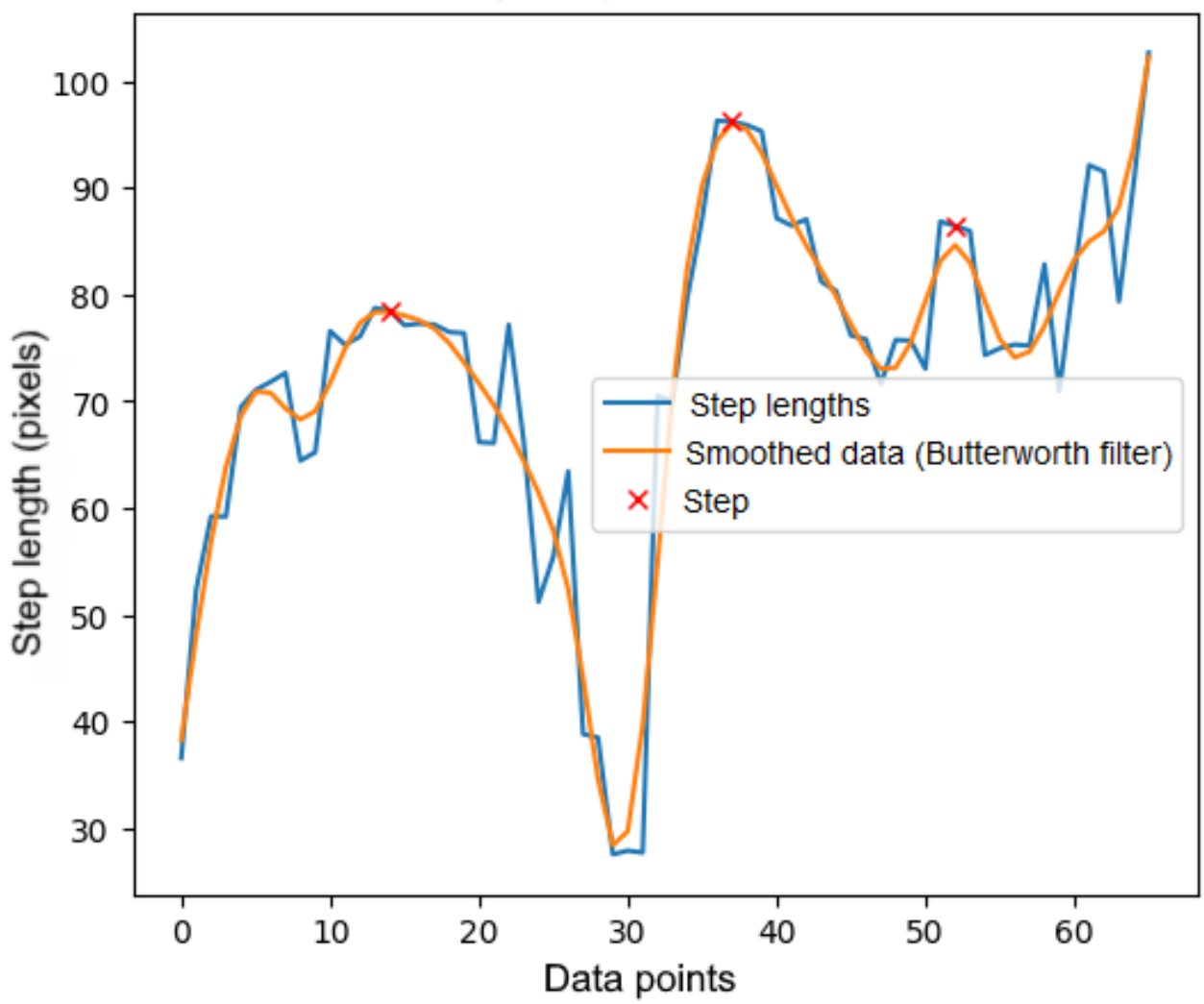
where $y(t)$ is the filtered signal; $x(t)$ is the input signal; fc is the cutoff frequency; f is the frequency of the signal; and n is the order of the filter.

In this study, an order of 4 was used for the Butterworth filters because it provides a smooth yet sufficiently steep roll-off in the transition band. While higher-order filters provide sharper roll-offs, they also require more computational resources. Therefore, a fourth-order filter appeared to be a reasonable compromise, being computationally efficient while providing adequate filtering. On the other hand, the peak detection equation is given by equation 2:

where $y(t)$ is the filtered signal; T is the set of all time points in the signal $y(t)$; t_i represents the time point where a peak occurs; and *height* is the threshold height for peak detection.

Only peaks that have an amplitude greater than or equal to the threshold, h, are detected and included in the output. In this study, the value for the height parameter was set to 100. An example of the graph for noise filtering with peaks is illustrated in Figure 5. This example demonstrates that the application of the Butterworth filter smooths the signal and yields an accurate step count during body turning.

Figure 5. An example of a graph processed with a Butterworth filter, showing peaks.



PD Recognition Using Model-Based Features

Certain gait features can help identify abnormalities or changes in walking patterns, which are particularly useful for diagnosing and monitoring conditions such as PD. By focusing on the body turning period, we aimed to capture the most challenging part of the gait cycle, where gait variabilities for PD are more easily observed. Turning involves complex motor coordination and balance adjustments, making it a difficult movement. This complexity can accentuate subtle abnormalities in gait patterns that might not be as apparent during straight walking. Therefore, analyzing gait features during turns provides a more sensitive and accurate assessment of PD-related gait abnormalities, enhancing the effectiveness of early diagnosis and monitoring. The gait features used in this study are summarized in Table 2.

Table 2. Gait analysis feature definitions and formulas.

Feature	Definition	Formula
Shoulder distance	Average horizontal distance between the left and right shoulders	mean(Shoulder_Distance($LShoulder_x$,$LShoulder_y$,$RShoulder_x$,$LShoulder_y$))
Step length	Average distance between the left ankle and right ankle	mean(Step_Length($LAnkle_x$,$LAnkle_y$,$RAnkle_x$,$RAnkle_y$))
Stride length	Average distance covered in 1 full stride, which consists of 2 steps (1 by each foot)	mean(Stride_Length($RAnkle_x$,$RAnkle_y$))
Angle of both knees	Degrees at the knee joint by considering the vectors formed by the knee to hip and knee to ankle points	Θ_{Knee}(Hip_x,Hip_y,$Knee_x$,$Knee_y$,$Ankle_x$,$Ankle_y$)
Angle of both hips	Degrees at the hip joint by considering the vectors formed by the hip to knee and hip to shoulder midpoint	Θ_{Hip} ($Shoulder_{mid}$,$Shoulder_{midy}$, Hip_x. Hip_y, $Knee_x$, $Knee_y$)
Symmetrical leg	Degrees for each ankle by considering the vectors formed by the knee to hip midpoint and knee midpoint to hip midpoint	Θ_{Leg}($Knee_{mid}$,$Knee_{midy}$, Hip_{mid} . Hip_{midy}, $Knee_x$, $Knee_y$)
Symmetrical arm	Degrees for each arm by considering the vectors formed by the arm to shoulder midpoint and arm midpoint to shoulder midpoint	Θ_{Arm}($Shoulder_{mid}$,$Shoulder_{midy}$, $Elbow_{mid}$. $Elbow_{midy}$)
Trunk Angle 1 (vertical)	Degrees by considering the vertical reference vector and the vector from the hip midpoint to the nose	Θ_{Trunk1}(Hip_{mid} ,Hip_{midy}, $Nose_x$. $Nose_y$)
Trunk angle 2 (horizontal)	Degrees by considering the vector from the hip midpoint to the shoulder midpoint and a horizontal reference vector	Θ_{Trunk2}($Shoulder_{mid}$,$Shoulder_{midy}$, Hip_{mid} ,Hip_{midy})
Shank angle	Degrees at the hip joint by considering the vectors formed by the left ankle to knee midpoint and right ankle to knee midpoint	Θ_{Shank}($Ankle_x$,$Ankle_y$,$Shoulder_{mid}$,$Shoulder_{midy}$)

Body Turning Duration Calculation

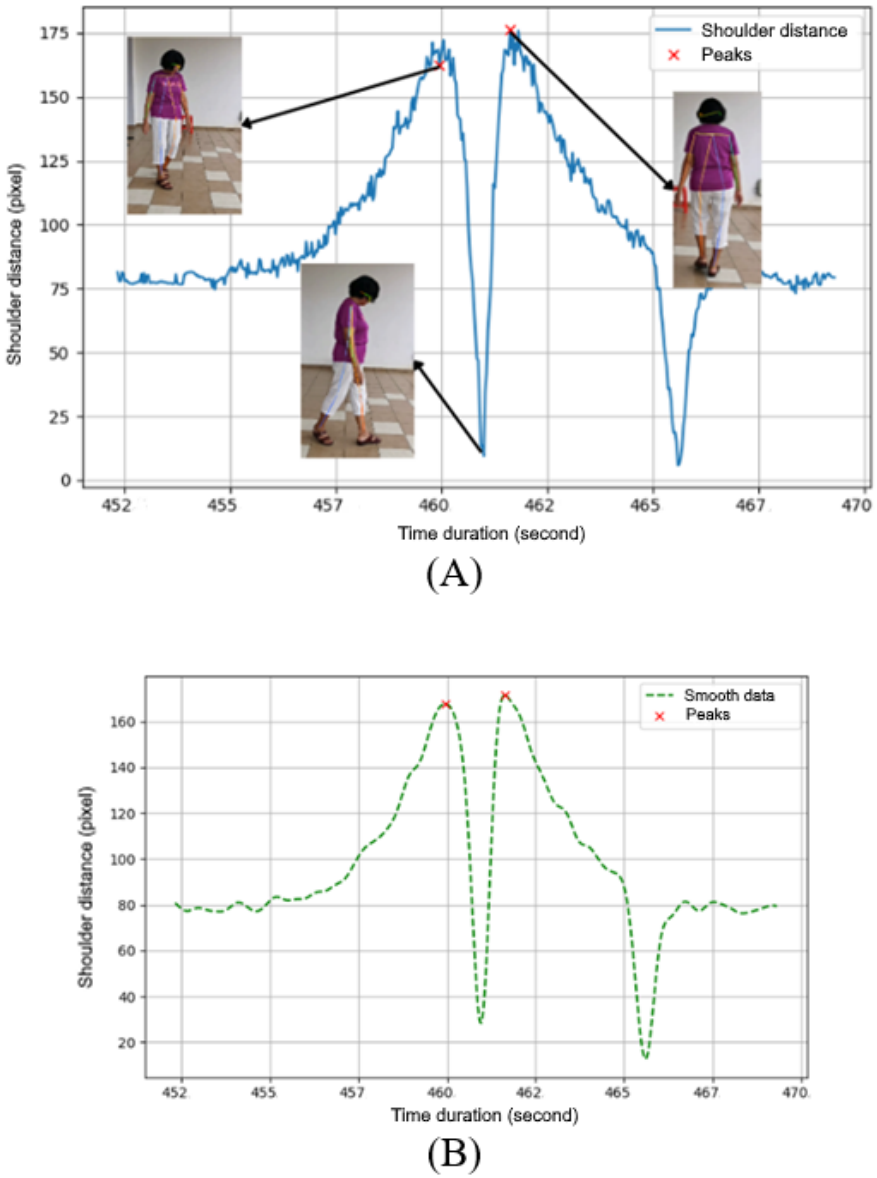
Figure 6 presents 2 graphs related to the analysis of shoulder distance during a body turn. The left graph (Figure 6A) displays the original, unprocessed data, illustrating the raw measurements of shoulder distance over time. The fluctuations observed may correspond to the natural movement variations during a turn. In contrast, the right graph (Figure 6B) exhibits data that have been smoothed using a Butterworth filter, a technique that mitigates short-term fluctuations and reveals the underlying pattern of movement more clearly.

The smoothed data enables effective determination of the peaks of the signal, which identifies the significant turning points. These peaks are highlighted by the red "X" marks in the figure and represent moments where the shoulder distance reaches its maximum, indicating a complete turn or a change in direction. By focusing on the 2 highest peaks, the graph underscores the most substantial turning events, thereby minimizing the potential for misinterpreting minor variations as significant movements.

In the context of shoulder distance, *LShoulder* and *RShoulder* would correspond to the left and right shoulders, with their respective *x* and *y* coordinates. The shoulder distance, *Shoulder_Distance*, which is a measure of how far apart the shoulders are, is calculated as follows (equation 3):

The indexes between these 2 peaks are then taken to determine the exact time the turn occurred. The difference between these peak times is calculated to find the duration between consecutive turning points. Let $T_{peak,i}$ be the time of the *i*th peak and *n* is the total number of peaks detected. Equation 4 calculates the total time duration between the first and the last peak by adding up the durations between each pair of consecutive peaks,

Figure 6. Shoulder distance analysis with peak detection: (A) graph before smoothing and (B) graph after smoothing.



Ethical Considerations

Ethical approval was granted by the Research Ethics Committee Multimedia University (approval number EA0422022). Informed consent was obtained from all the participants. A statement about this is given in the Methods section. Identifiable features of research participants are not visible in the manuscript. No compensation was provided to the participants.

Results

Overview

This section presents the experiment results and discussion for the gait features extracted from the self-collected dataset, focusing on the effectiveness of these features in distinguishing between normal and PD-affected gait patterns using a SVM classifier.

For the experiments, principal component analysis was applied to reduce the dimension of the dataset. The number of components for principal component analysis was set to 8, corresponding to the number of participants with PD. This step helps in capturing the most significant features while reducing computational complexity.

The dataset was additionally divided into a 70:30 ratio. This indicates that the model was trained using 70% of the data, with the remaining 30% set aside for testing and performance evaluation. By splitting the data this way, a significant portion of the data were used to train the model while maintaining enough for a thorough assessment. In total, 10 trials of the experiments were performed, and the average results were recorded.

Participant Details

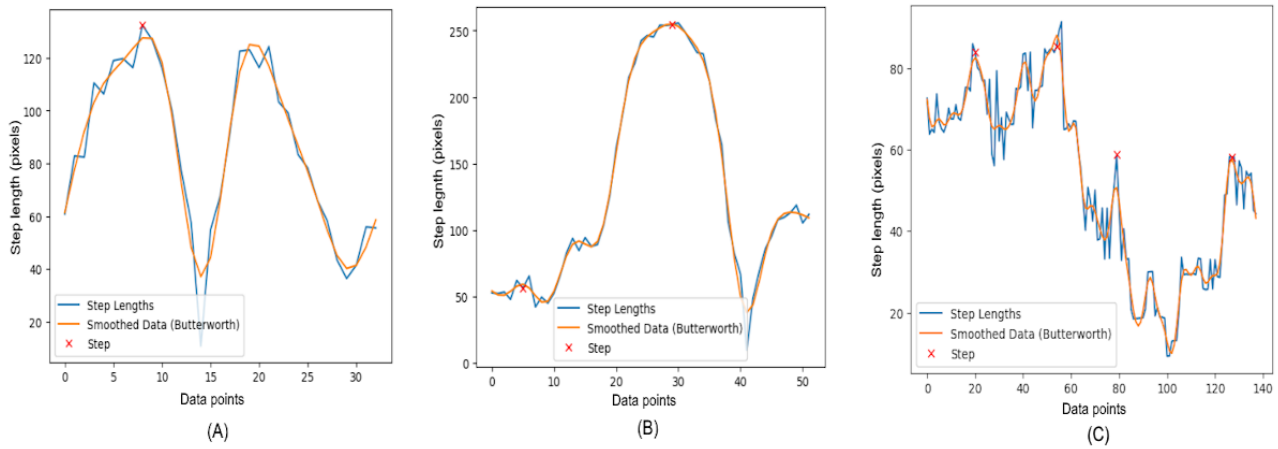
There were 3 groups of participants in the self-collected dataset, namely, patients with PD, older adults, and adolescents. A significant portion of the participants (8/28, 29%) were diagnosed with PD, while the remaining participants were healthy (20/28, 71%). Most of the participants were Chinese (24/28, 85%), with a smaller representation of Indian (3/28, 11%) and Malay (1/28, 4%) individuals. There was a higher

proportion of male participants (21/28, 75%) compared to female participants (7/28, 25%). Finally, most participants fell within the age range of 60 to 69 years (14/28, 50%), followed by the age range of 20 to 29 years (9/28, 32%), and smaller proportions in other age ranges. These results collectively illustrate the diversity and focus of the study population.

Total Steps During Body Turning

Figure 7 presents 3-line graphs that correspond to the step lengths for different demographic groups: younger individuals, older individuals, and individuals with PD. The step lengths are measured during a turning sequence.

Figure 7. A comparative step analysis by step length between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



The graph of younger individuals (Figure 7A) demonstrates a more variable stepping pattern with pronounced peaks, suggesting agility and a higher range of motion (ROM) during turns. The steps are uneven in length, but the peaks return to baseline quickly, indicating swift changes in direction and a potentially more dynamic gait.

In contrast, the graph of older individuals (Figure 7B) has fewer, more rounded peaks, indicative of a more cautious or steady turning strategy. The step lengths are generally more uniform, with smoother transitions between steps. This could reflect a more deliberate and potentially less stable gait, as is often seen with aging.

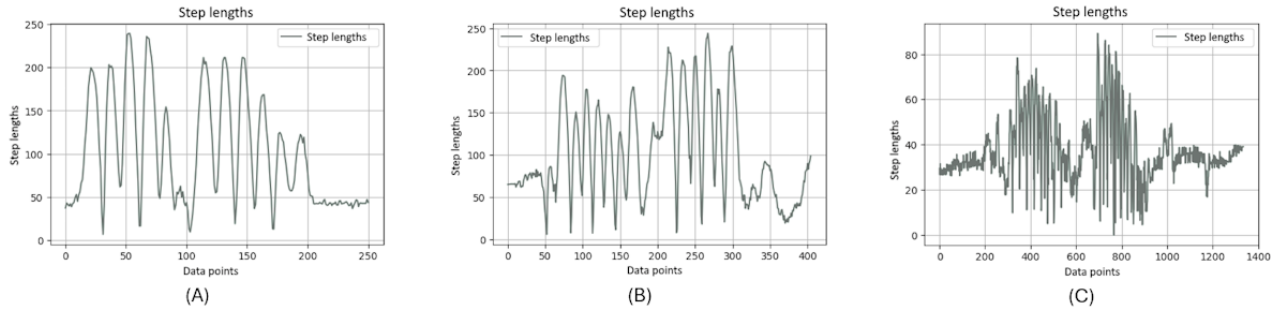
The graph of individuals with PD (Figure 7C) demonstrates a significantly different pattern, with smaller step lengths and a higher frequency of steps, which may point to the short, shuffling steps often associated with PD. The peaks are less pronounced and more erratic, highlighting the challenges individuals with PD may face in maintaining a regular stepping pattern.

Step Length

Step length refers to the distance between the left and right ankle for each step. It is determined by the coordinates of the left ankle ($LAnkle_x$, $LAnkle_y$) and the right ankle ($RAnkle_x$, $RAnkle_y$) at a particular instance in time. The formula to calculate step length is provided in equation 5. This formula gives the straight-line distance between 2 points in a 2D space, such as a video frame or a motion capture system's coordinate plane.

In Figure 8, the graphs illustrate the differences in gait patterns across 3 demographic groups during a TUG test. Each graph plots the variation in step lengths, which reflect the distance between the left and right ankle during individual steps.

Figure 8. A comparative step length analysis between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



The graph of younger individuals (Figure 8A) is characterized by high peaks and deep valleys, indicating significant variability in step lengths. This could suggest a dynamic and robust gait, with the ability to take both long and short steps, perhaps adjusting speed or direction more frequently.

In contrast, the graph of older individuals (Figure 8B) shows a moderate level of variability with somewhat rounded peaks.

The reduced height of the peaks compared to the young group suggests shorter steps on average, which may be a sign of a more cautious approach to movement, possibly due to decreased mobility or balance concerns that come with age.

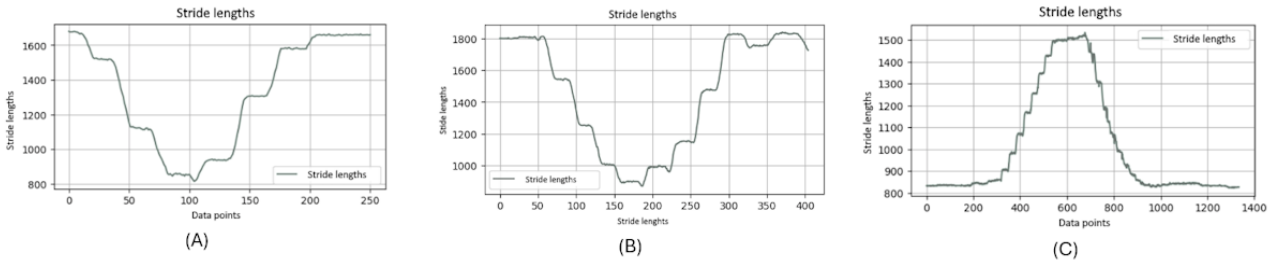
The graph of individuals with PD (Figure 8C) differs markedly from the other 2. It has a much higher frequency of smaller fluctuations, and the overall range of step lengths is noticeably lower. This pattern indicates the short, shuffling steps that are often observed in individuals with PD, reflecting the challenges they face with gait initiation and continuation.

Stride Length

Stride length is typically defined as the distance covered in one full stride, consisting of 2 steps (one by each foot). The equation to calculate the stride length for each frame is presented in equation 6. It calculates the stride length as the Euclidean distance from a point (possibly the origin) to the coordinates of the right ankle for each frame.

Figure 9 shows the stride length of the right ankle throughout a single TUG video. The graph exhibits the dynamic nature of stride lengths, with the vertical axis representing the stride length and the horizontal axis corresponding to data points that likely represent sequential frames of the video.

Figure 9. A comparative stride length analysis between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



It is noteworthy that this graph clearly distinguishes between 2 types of motion: the rising and falling edges indicate the leg in motion, not in contact with the ground, and the flat regions signify the stance phase where the leg is stationary on the floor, awaiting the next step. The consistency in the flat regions indicates moments when the leg is at rest between strides.

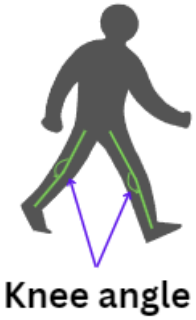
It is evident that the graphs of younger and older individuals (Figures 9A and 9B) display similar frequencies in stride patterns, whereas the graph of the individuals with PD (Figure 9C) significantly diverges. It illustrates a strikingly different frequency in stride lengths, standing out from the more consistent rhythmic patterns observed in the younger- and older-individual groups.

Moreover, from these charts, it is obvious that when younger and older people complete a stride, their feet leave the ground to a greater extent, because a normal stride will take a relatively large range. On the contrary, when patients with PD complete a stride, their feet leave the ground to a relatively smaller extent. The amplitude of the ground is relatively short, which can be considered a symptom of PD.

Angle of Both Knees

Analyzing the positions of the hip, knee, and ankle for each leg to determine the angles at both knees can assist in recognizing PD (Figure 10).

Figure 10. An illustration of the knee angle position.

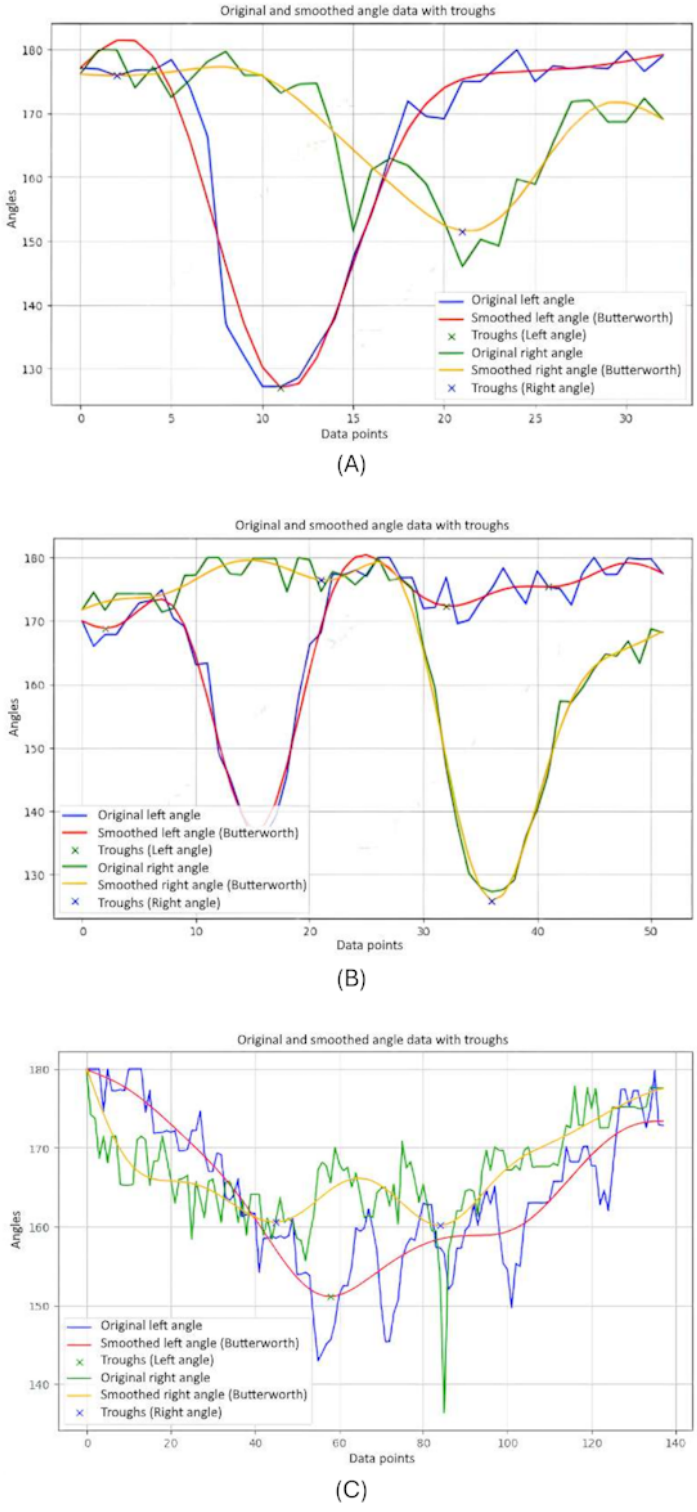


The knee angles provide crucial insights into the degree of bending or extension at the knee joints, which are essential for understanding gait patterns. PD often affects movement and posture, leading to distinctive gait abnormalities. By quantifying these knee angles, the calculation helps in detecting and assessing these gait characteristics, aiding in the identification and monitoring of PD symptoms.

Given the knee, hip, and ankle joint positions, the knee angle and the Θ_{knee} can be computed as follows (equation 7):

Figure 11 presents 3-line graphs comparing the knee joint angles of the 3 groups: younger individuals, older individuals, and individuals with PD. The graph of the younger individuals group (Figure 11A) illustrates smooth and regular oscillations in knee joint angles for both the left and right legs. As they turn their body, the angles show significant flexion and extension, reflecting the youthful ability to execute the turn in a single, fluid step. The peaks and troughs are well-defined, indicating robust and agile movements typical of healthy, young individuals. This pattern highlights their efficient and coordinated gait.

Figure 11. A comparative analysis of the knee angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



In the older-individual group graph (Figure 11B), the knee joint angles are more subdued, with 2 distinct curves for each turn. This indicates that older individuals take 2 steps to complete the body turn, reflecting a more cautious and segmented approach. The movement patterns are slower and less dynamic compared to the younger individual group, with a reduced ROM.

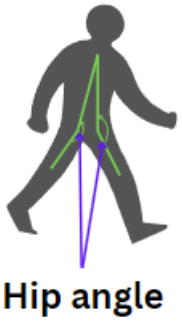
The graph for the group of individuals with PD (Figure 11C) shows highly irregular and erratic knee joint angles. The curves exhibit frequent sharp peaks and drops, indicating inconsistent and disrupted movement as individuals with PD attempt to turn their bodies. These fluctuations reflect the challenges faced by patients with PD, including tremors, rigidity, and difficulty maintaining smooth and coordinated movements. The graph captures their struggle to control knee flexion and extension, leading to a more disordered and interrupted gait pattern during the turning motion.

These graphs highlight clear differences in turning movements: young individuals exhibit fluid, single step turns; older individuals use a careful 2-step approach; and patients with PD show irregular, disrupted turning patterns.

Angle of Both Hips

The ROM angle while turning the body is used to determine the angle of both hips (Figure 12). The positions of the hip, knee, and shoulder are used to find the angle formed at the hip joint. This measurement is crucial for assessing how the hip moves during turns, providing insights into the flexibility and coordination of the lower body. Understanding the hip ROM is especially valuable for identifying movement patterns in different populations, such as detecting mobility issues or assessing gait in individuals with PD.

Figure 12. An illustration of the hip angle position.



The hip angle is calculated as follows (equation 8):

The graphs showing the cosine of the hip angle are shown in Figure 12. The figure shows 3-line graphs comparing hip angles during body turns for the 3 groups. Each graph represents the hip angles for both the left and right sides, with lines depicting both the original and smoothed data.

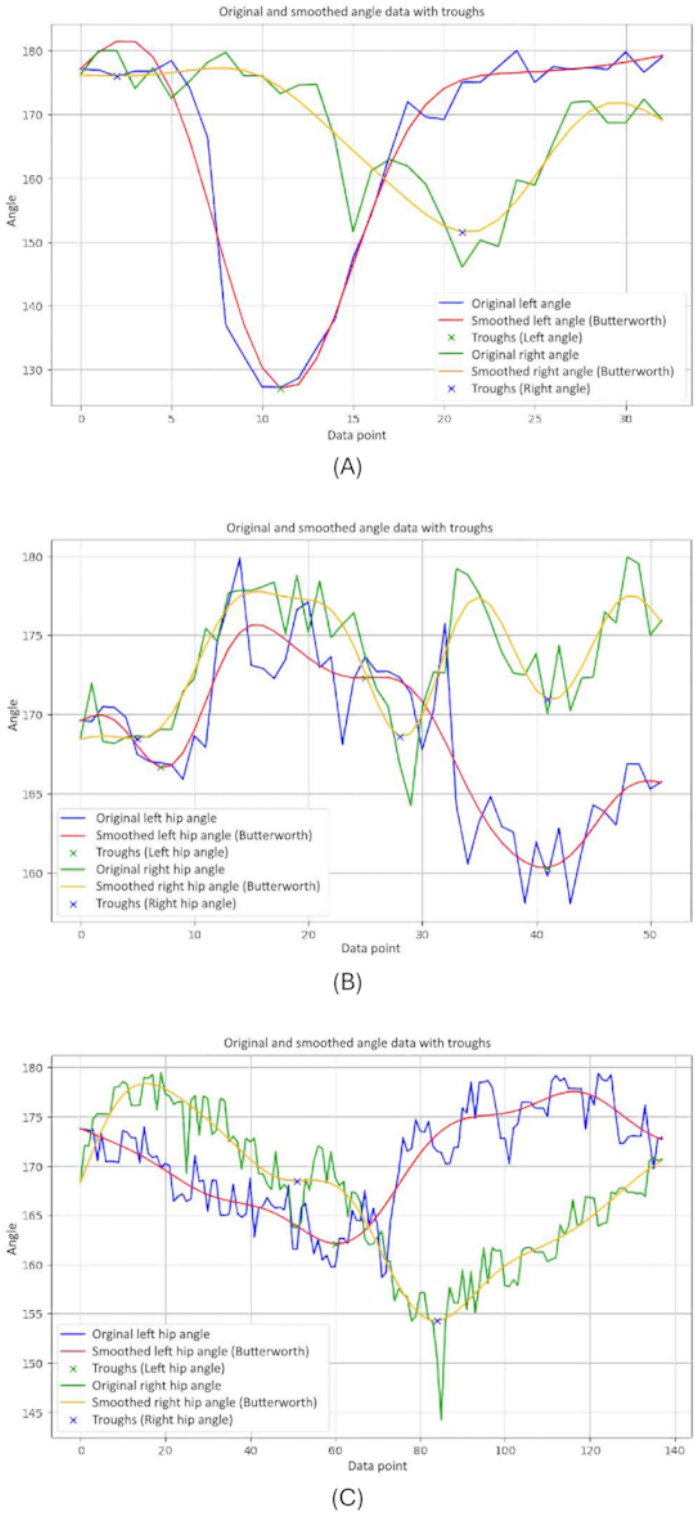
The graph of the younger individuals group (Figure 12A) indicates relatively stable and moderate fluctuations in hip angles during body turns. Both hips show consistent movement with shallow peaks and troughs, reflecting balanced and controlled motion. Younger individuals maintain a steady range of hip angles, suggesting efficient and coordinated hip movement when turning.

In the older-individual group graph (Figure 12B), the hip angles display more pronounced and variable fluctuations compared to the young group. The curves reveal that older individuals experience larger and more variable hip movements during turns. This increased variability suggests that older people adjust their hip movements significantly to maintain balance and stability, resulting in less smooth and more fluctuating hip angles.

The group of individuals with PD (Figure 12C) shows significant irregularity and instability in hip angles. The curves exhibit erratic and sharp changes, highlighting the difficulty individuals with PD face in maintaining consistent hip movement during turns. The frequent and abrupt fluctuations in hip angles are characteristic of the disease, where symptoms such as rigidity and tremors disrupt smooth and coordinated movements.

Figure 13 illustrates clear differences in hip angle dynamics during turns among the 3 groups. Younger individuals (Figure 13A) have controlled and consistent hip movements, older individuals (Figure 13B) show more variability and larger fluctuations in hip angles, and those with PD (Figure 13C) experience erratic and unstable hip movements. These patterns reflect how age and neurological conditions affect the coordination and efficiency of hip movement during body turns.

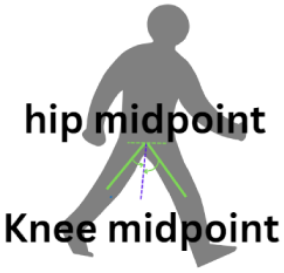
Figure 13. A comparative analysis of the hip angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Leg Symmetry

The symmetry of leg movements is calculated using the angles at the knee joints (Figure 14). The goal was to measure and evaluate the balance and symmetry in the leg movements, which is crucial for understanding gait and stability, especially during turning motions. This information is particularly useful for analyzing the movement patterns in different populations, such as detecting gait irregularities in individuals with conditions such as PD.

Figure 14. An illustration of the symmetrical leg position.



The leg angle is computed as follows (equations 9-11):

Next, the angles of the left and right legs are compared:

The leg symmetry is then assessed as:

Figure 15 displays 3-line graphs comparing the leg symmetry during body turns for the 3 participant groups. Each graph plots the symmetrical leg movements, showing the original and smoothed angles of the left and right legs as they turn.

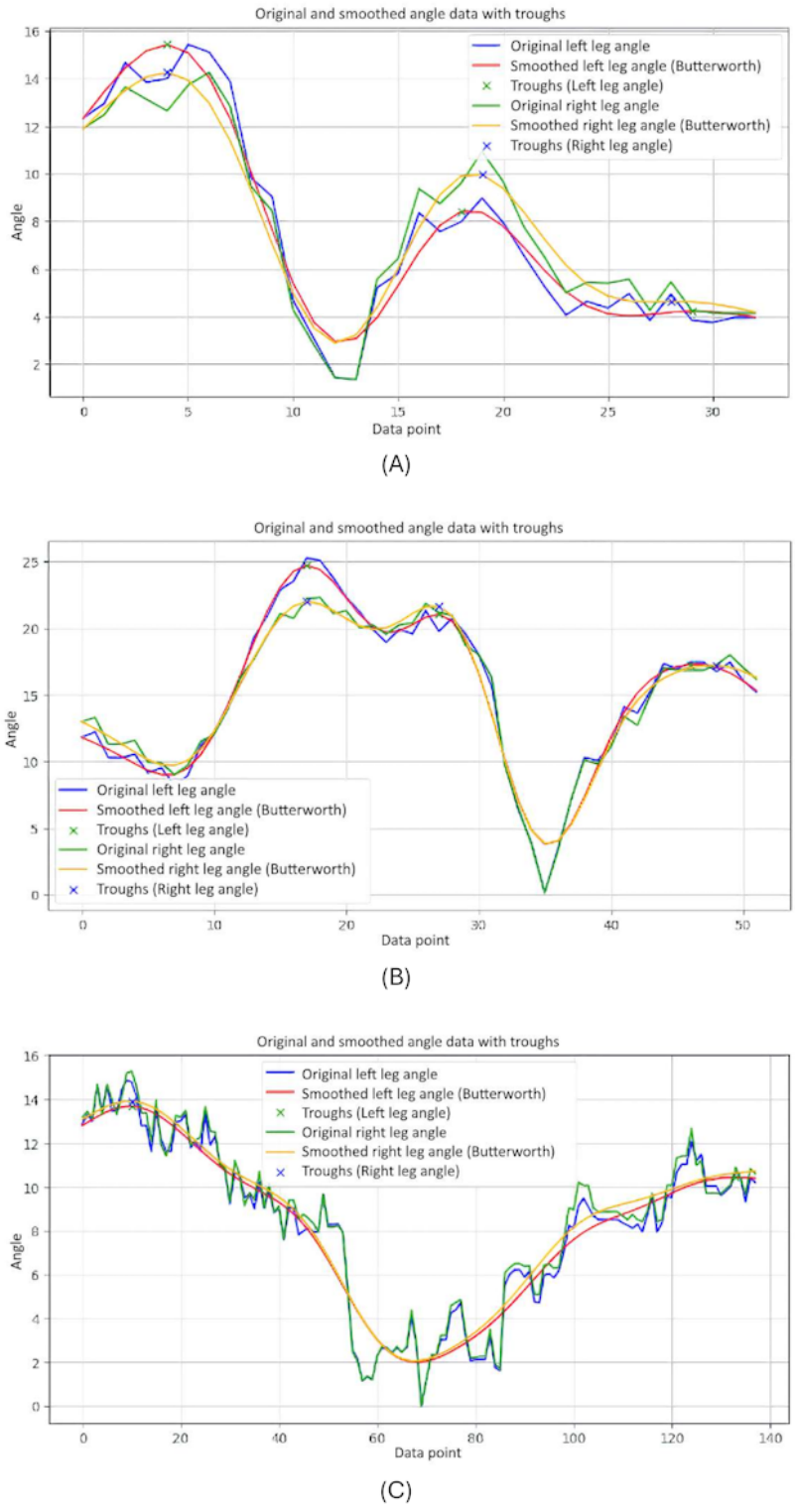
The graph of the younger-individual group (Figure 15A) shows closely aligned curves for the left and right leg angles, indicating high symmetry in their leg movements during turns. Both legs exhibit similar patterns with minimal deviation between them. This close alignment reflects the balanced and coordinated movements typical of young, healthy individuals, suggesting that they maintain a consistent and symmetrical gait.

The older-individual group graph (Figure 15B) reveals more noticeable fluctuations between the left and right leg angles. While the overall patterns still follow a similar trajectory.

On the other hand, the graph of individuals with PD (Figure 15C) shows significant irregularities and less alignment between the left and right leg angles. The curves are erratic and display frequent sharp deviations, reflecting low symmetry in their leg movements. This lack of alignment is characteristic of PD, where symptoms such as tremors and rigidity cause disrupted and uncoordinated movements.

This comparison highlights how age, neurological conditions, and health status impact the ability to maintain balanced and symmetrical gait during turning motions.

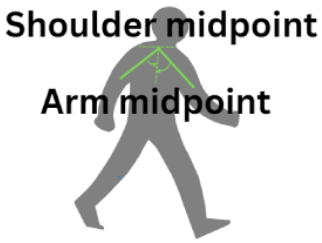
Figure 15. A comparative analysis of the symmetrical leg between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Arm Symmetry

The symmetry of arm movements was analyzed using the angles at the elbows to the shoulders (Figure 16). The purpose was to measure how balanced and coordinated the arm movements are, especially during activities such as turning the body. This assessment was important for understanding the ROM and the symmetry in arm movements, which can provide insights into overall body coordination and detect possible imbalances or movement disorders.

Figure 16. An illustration of the symmetrical arm position.



The calculation for arm symmetry was similar to that of finding the symmetry in legs (equation 12):

Figure 17 presents 3-line graphs comparing the arm symmetry during body movements for 3 groups. Each graph plots the angles of the left and right arms over time, showing both the original and smoothed data.

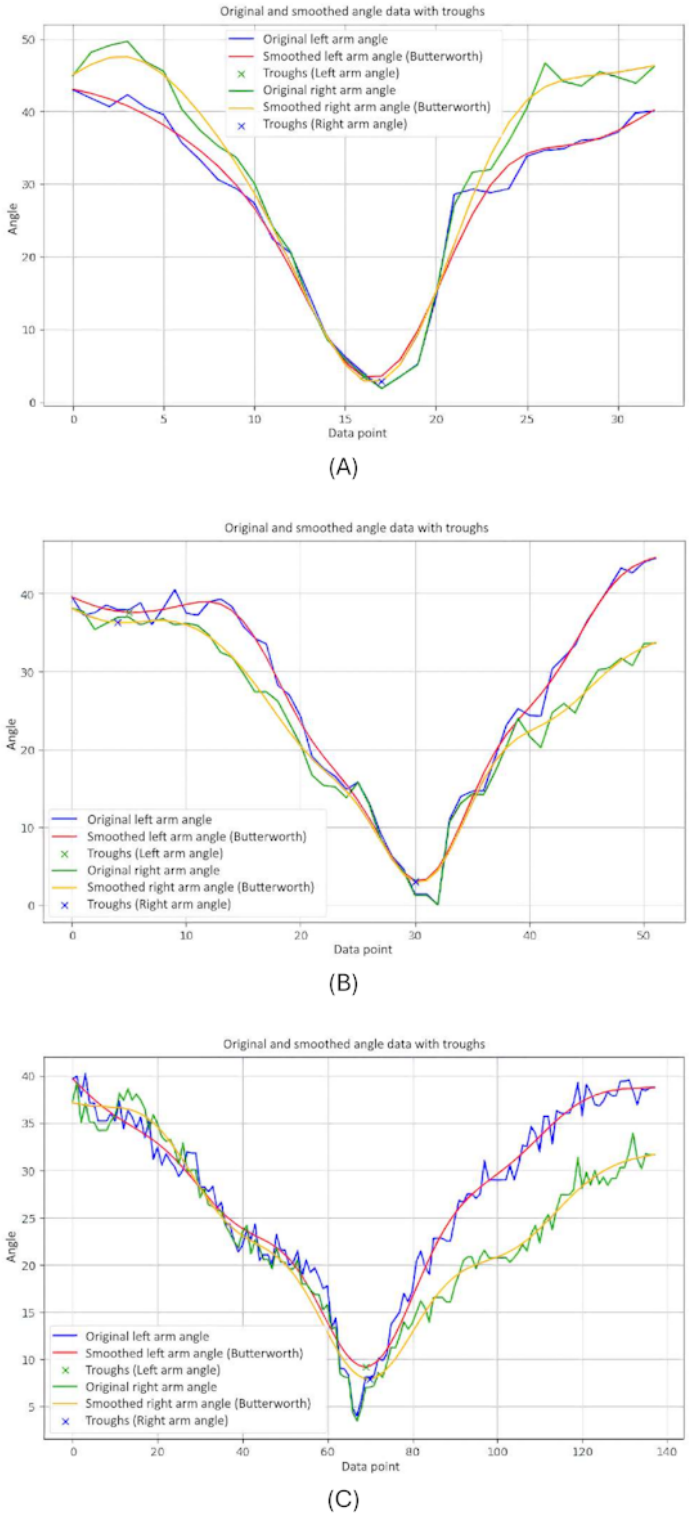
The graph of the younger-individual group (Figure 17A) displays closely aligned curves for the left and right arm angles. Both arms move in a highly synchronized manner, with minimal differences between the left and right angles throughout the motion.

The older-individual group graph (Figure 17B) shows a slightly more varied pattern compared with the young group. While the overall movements of the left and right arms remain relatively synchronized, there are noticeable differences in the magnitude and timing of the angles.

The graph of individuals with PD (Figure 17C) exhibits significant irregularities and discrepancies between the left and right arm angles. The curves are erratic and often diverge sharply, indicating substantial asymmetry in arm movements.

We observed that younger individuals demonstrate highly synchronized and symmetrical arm movements, older individuals show moderate symmetry with some variability, and those with PD experience significant asymmetry and irregularity in their arm movements. These patterns underscore how age and neurological conditions impact the coordination and balance of arm movements during body turns.

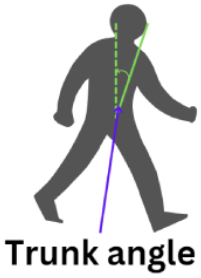
Figure 17. A comparative analysis of the symmetrical arms between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Trunk Angle 1 (Vertical)

The relative positions of the hips and nose are used to calculate the trunk angle (Figure 18). The trunk angle helps to understand how much the upper body tilts or bends relative to the lower body, especially during activities such as turning. This angle helps in assessing posture and balance, providing insights into the coordination and alignment of the trunk with the hips during movement. It is particularly useful in evaluating movement patterns and detecting postural deviations in various populations, including those with movement disorders such as PD.

Figure 18. An illustration of the trunk angle 1 vertical position.



The calculation of the trunk angle 1 is given as follows (equation 13):

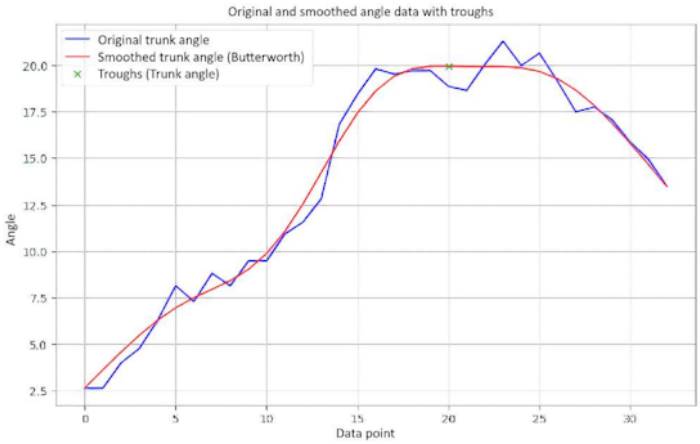
Figure 19 displays 3 graphs comparing trunk angles relative to the vertical axis for 3 groups. Each graph shows the trunk angle over a series of data points during body movements, with both original and smoothed data represented. The trunk angle is used to observe the severity of hunchback, providing insights into the degree of forward bending or curvature of the upper body.

The graph of the younger-individual group (Figure 19A) exhibits a consistent and smooth progression of trunk angles. The trunk angle gradually increases and decreases within a narrow range, reflecting a well-coordinated and balanced posture with minimal forward bending. Younger individuals maintain a steady alignment with minimal deviations, indicating a lack of significant hunchback severity.

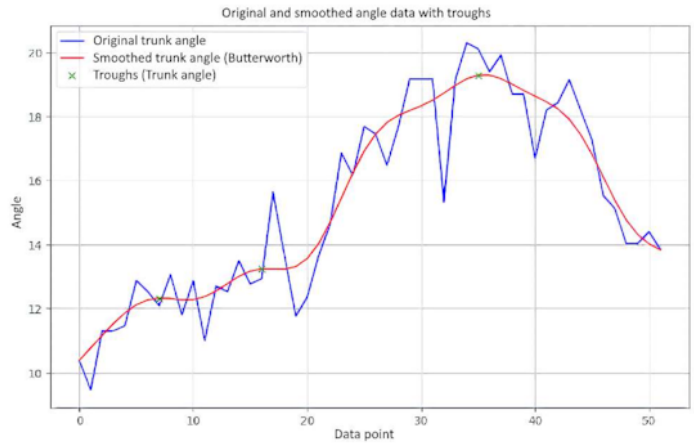
The older-individual group graph (Figure 19B) shows more variability in trunk angles compared to the young group. The curves display noticeable fluctuations and less consistency, with several abrupt changes in angle. These variations suggest that older individuals experience more difficulty in maintaining a stable trunk posture, leading to less smooth and more erratic movements. The increased trunk angles and variability may indicate a greater tendency toward forward bending, suggesting a moderate severity of hunchback as they struggle to maintain an upright posture.

The graph of individuals with PD (Figure 19C) reveals significant irregularities and instability in trunk angles. The curves are highly erratic, with frequent sharp changes and a wide range of deviations from the vertical alignment. These abrupt shifts indicate substantial difficulties in maintaining consistent trunk posture. Individuals with PD struggle with controlling their trunk movements, leading to frequent tilting and misalignment relative to the vertical axis. The pronounced forward bending and high variability in trunk angles reflect severe hunchback, characteristic of the disease's impact on posture and movement control.

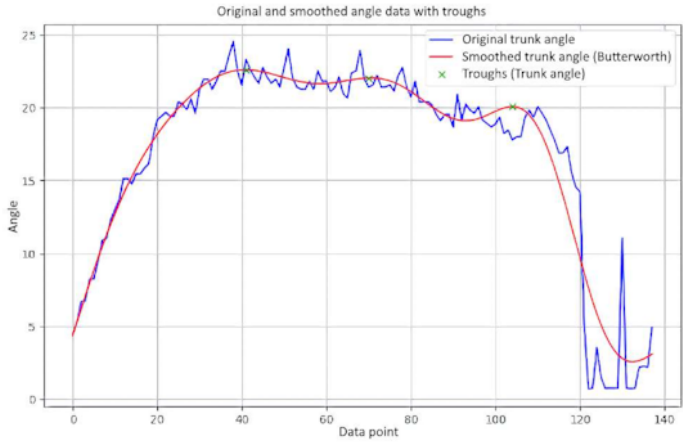
Figure 19. A comparative analysis of the trunk angle 1 between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



(A)



(B)

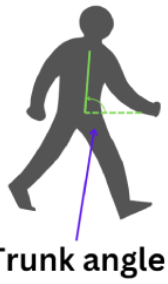


(C)

Trunk Angle 2 (Horizontal)

Trunk angle 2 analyzes the positions of the shoulders and hips (Figure 20). This angle is crucial for understanding the alignment of the trunk relative to the lower body, particularly during movements such as turning. By evaluating the trunk angle, we can assess the severity of hunchback or forward bending, which provides valuable insights into posture and stability.

Figure 20. An illustration of the trunk angle 2 horizontal position.



The calculation for trunk angle 2 is similar to that for trunk angle 1. The only difference is that as trunk angle 2 is horizontal, the default value must be set at the *x* coordinate instead of the keypoint. The formula to calculate trunk angle 2 is given as follows (equation 14):

Figure 21 shows graphs comparing trunk angles relative to the horizontal axis for 3 groups. These angles help assess the severity of forward bending or misalignment.

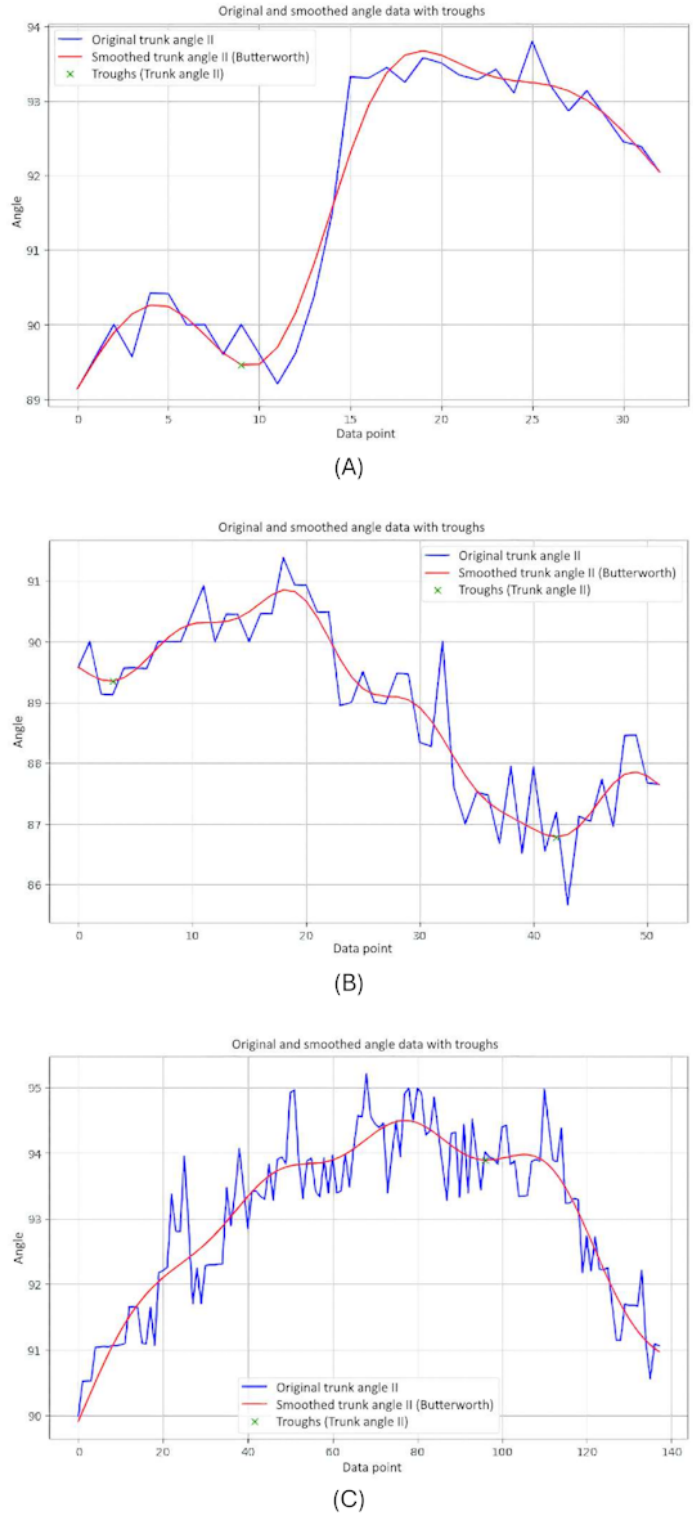
The graph of the younger-individual group (Figure 21A) exhibits stable and controlled trunk angles with minimal deviations, indicating well-balanced posture and low severity of hunchback.

The graph of the older individuals' group (Figure 21B) showed more variability and sudden changes in trunk angle, reflecting moderate forward lean and greater difficulty maintaining stable horizontal alignment, indicating moderate kyphosis severity.

The graph of the individuals with PD (Figure 21C) shows significant instability and frequent sharp changes in trunk angles, indicating severe difficulties in maintaining consistent trunk posture and severe hunchback severity due to the disease's impact on movement control.

Overall, younger individuals maintain a low severity of hunchback, older individuals exhibit moderate severity, and those with PD show severe hunchback, reflecting their challenges in maintaining horizontal alignment.

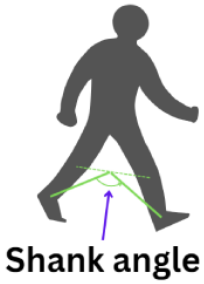
Figure 21. A comparative analysis of the trunk angle 2 horizontal between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Shank Angle

The shank angle computes the positions of the knees and ankles (Figure 22). This angle is useful to understand the alignment and movement of the lower leg, especially during activities such as turning. By evaluating the shank angle, we can assess the coordination and stability of the lower leg movements, providing valuable insights into gait patterns and detecting any abnormalities in movement, which is particularly useful for analyzing conditions such as PD.

Figure 22. An illustration of the shank angle position.



The formula to calculate shank angle is provided in equation 15:

Figure 23 presents the graphs comparing shank angles during body movements for 3 groups. Each graph shows the shank angle over a series of data points, with both original and smoothed data represented.

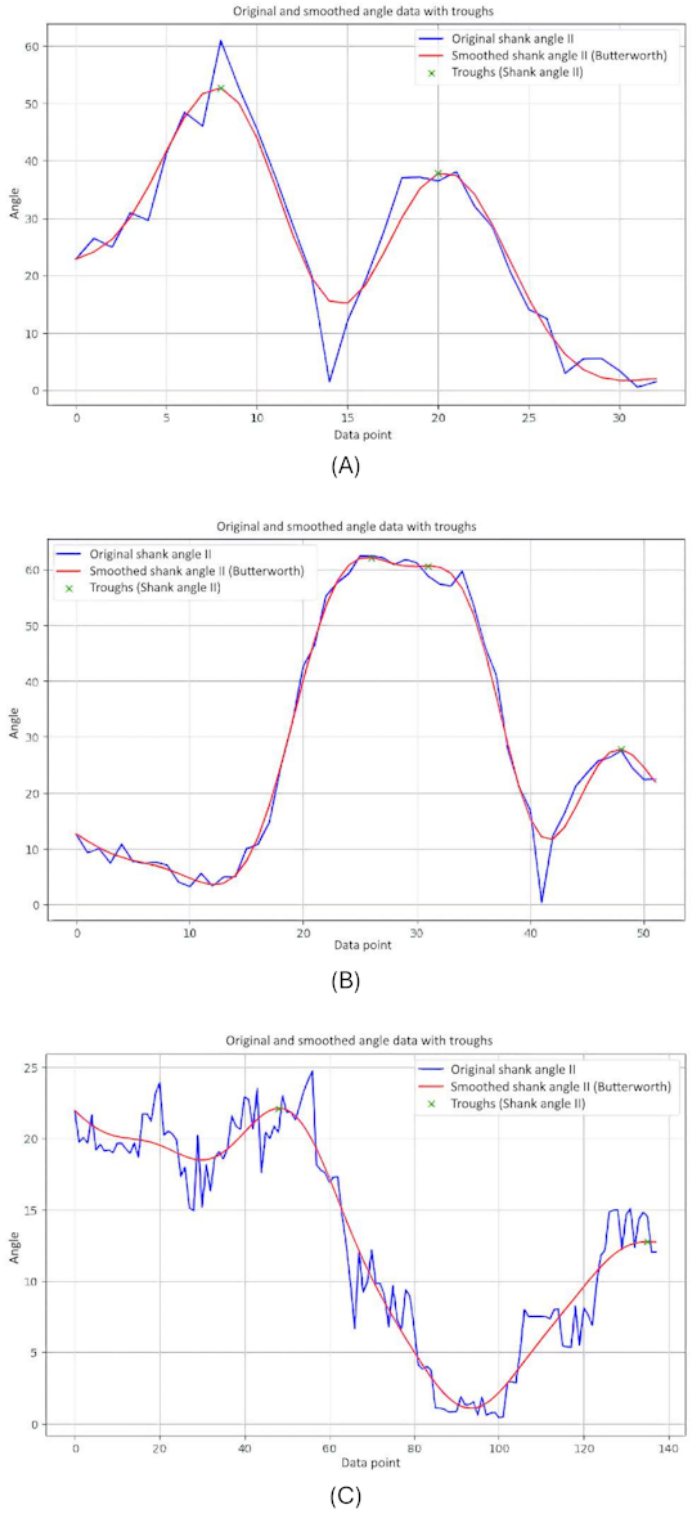
The graph of the younger-individual group (Figure 23A) demonstrates relatively stable and consistent shank angles. Both the original and smoothed curves align closely, indicating smooth and coordinated movements. The angles show moderate fluctuations within a narrow range, reflecting balanced and controlled lower leg movements typical of healthy young individuals.

The older-individual group graph (Figure 23B) shows shank angles that are stable and consistent, such as the younger individuals group. Both the original and smoothed curves align closely, indicating smooth and coordinated lower leg movements. The angles show moderate fluctuations within a narrow range, reflecting balanced and controlled movements. This suggests that older individuals perform nicely, maintaining stable shank angles comparable to the younger individuals group.

The graph of individuals with PD (Figure 23C) reveals significant irregularities and instability in shank angles. The curves are highly erratic, with frequent sharp changes and a wide range of deviations. These abrupt shifts indicate substantial difficulties in maintaining consistent lower leg movements. Individuals with PD struggle with controlling their shank angles, leading to frequent misalignment and instability. This pattern is characteristic of the disease, where symptoms such as rigidity and tremors disrupt smooth and coordinated movements.

Figure 23 highlights the differences in shank angle behavior among the 3 groups. Younger and older individuals exhibit stable and coordinated shank movements with minimal variability, indicating balanced and controlled lower leg movements. In contrast, those with PD display significant instability and erratic shank movements, underscoring the impact of the condition on lower leg control and stability.

Figure 23. A comparative analysis of the shank angle between different groups: (A) younger individuals, (B) older individuals, and (C) individuals with Parkinson disease.



Classification Report

A comprehensive evaluation of the SVM model is provided by the classification report for the test set, as shown in Table 3. The report includes the overall accuracy of the model, as well as each class's recall, precision, and F_1-score. Although the training accuracy was 0.95, the test accuracy was 0.89, indicating that the model fit the training data well with minimal overfitting. The classification report shows that class 0 (normal class) had high precision and recall, while class 1 (PD class) has lower recall. This disparity in recall may be due to the smaller sample size of class 1.

Table 3. Classification report for distinguishing normal and PD[a] classes using gait features (test accuracy: 0.89 and train accuracy: 0.95).

Classes or metrics	Precision	Recall	F_1-score	Support
0 (normal)	0.88	1.00	0.93	7.00
1 (PD)	1.00	0.50	0.67	2.00
Accuracy	—[b]	—	0.89	9.00
Macroaverage	0.94	0.75	0.80	9.00
Weighted average	0.90	0.89	0.87	9.00

[a]PD: Parkinson disease.

[b]Not available.

AUC and Loss Curve

Figure 24 provides specifics regarding the model's capacity to distinguish between the 2 classes. Moderate discrimination capability is indicated by the test set's AUC of 0.6429. An indicator of how well the model performs across various training data subsets is provided by the cross-validation scores. With a mean cross-validation score of 0.9, the model's cross-validation scores are (0.75, 1.00, 1.00, 0.75, 1.00). This shows that the model continues to work well when the data are folded in different ways.

Figure 24. The receiver operating characteristic (ROC) curve.

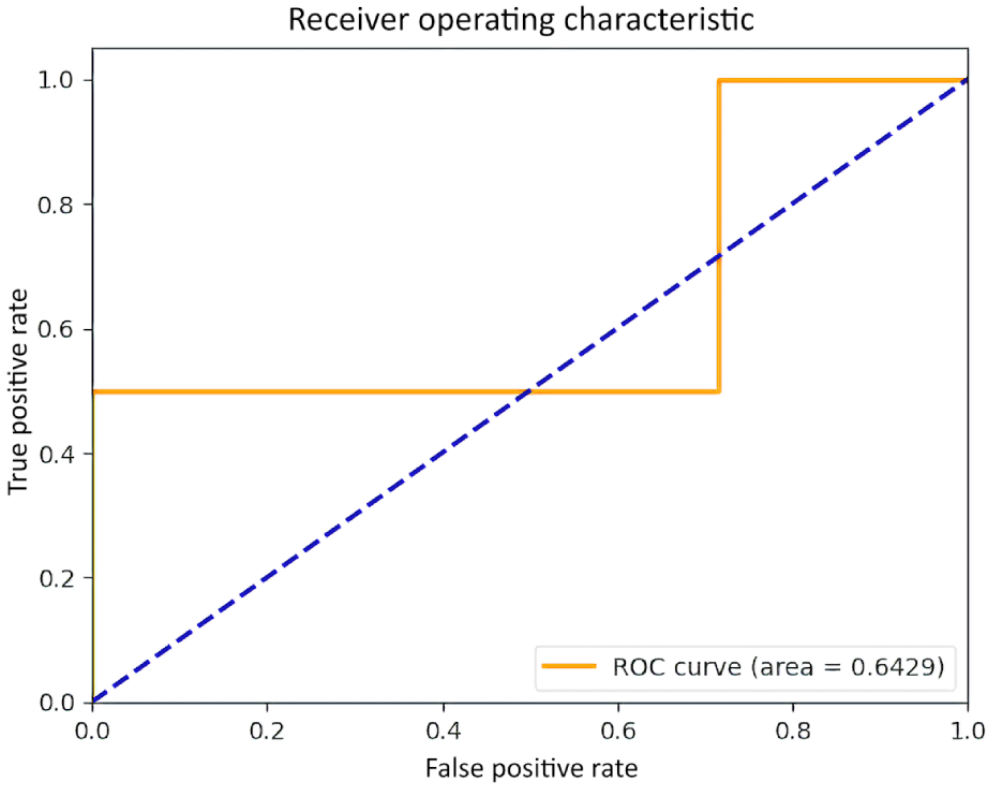
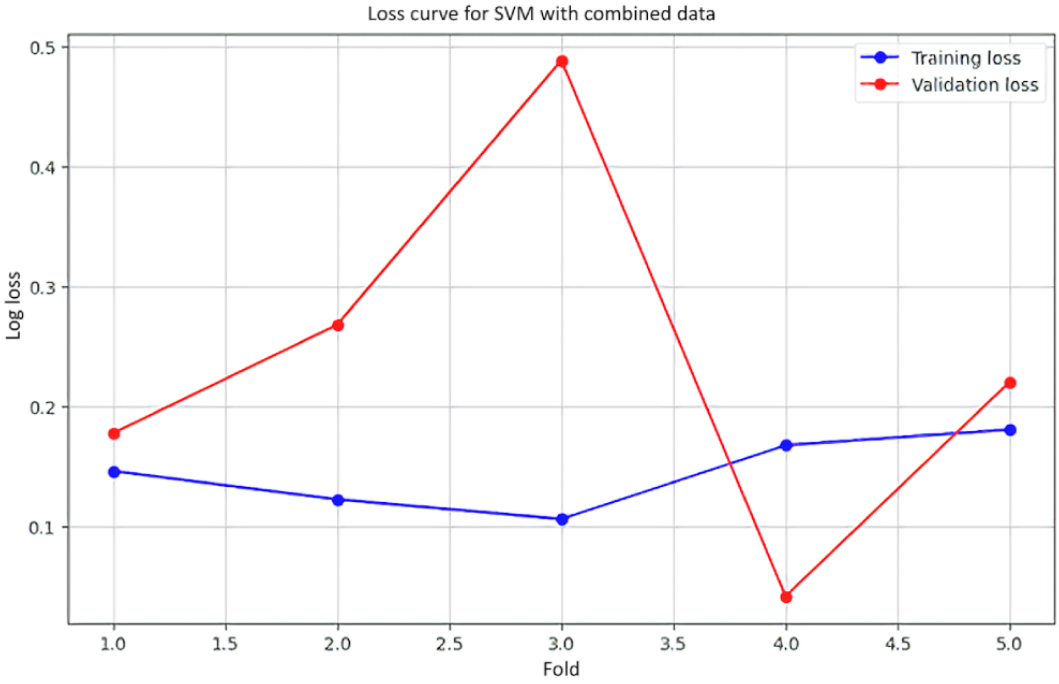


Figure 25 shows the training and validation loss curve, illustrating the log loss for both training and validation sets across various folds of cross-validation. The blue curve indicates the training loss, which stays relatively stable with minor fluctuations. In contrast, the red curve, representing the validation loss, exhibits significant fluctuations, peaking notably at fold 3. This implies that while the model fits the training data well, it experiences considerable variability in performance on unseen data, suggesting possible overfitting or sensitivity to specific data subsets. The difference between the training and validation loss also underscores the model's difficulty in generalizing from the training set to the validation set, highlighting the need for potential adjustments in model complexity or training strategy to achieve more consistent performance.

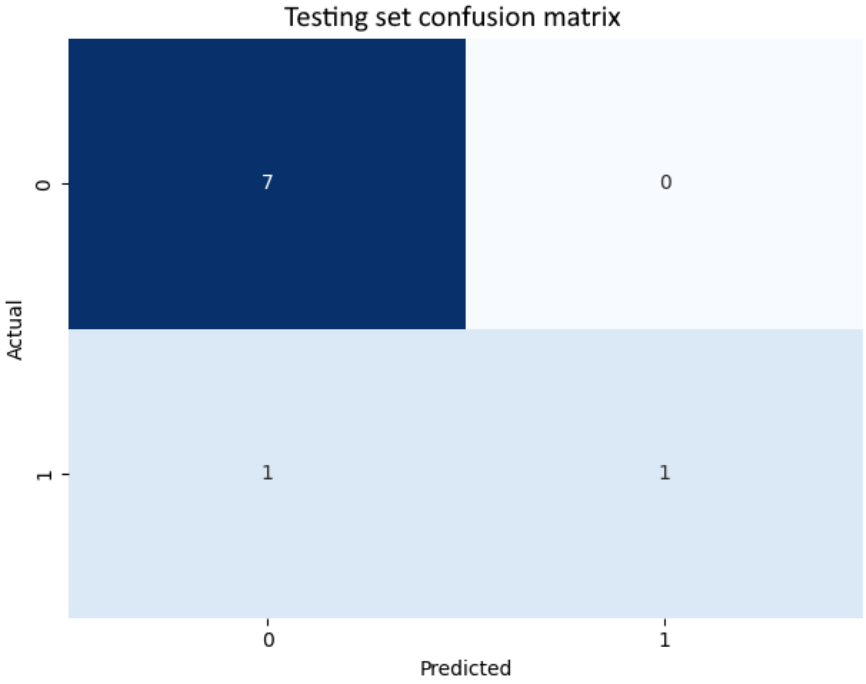
Figure 25. The training and validation loss curve for support vector machine (SVM) with combined data.



Confusion Matrix

The confusion matrix depicted in Figure 26 provides a thorough analysis of the predictions of the model on the test set. It reveals that while the model correctly predicted 7 out of 7 occurrences of class 0 (non-PD), it incorrectly classified 1 out of 2 instances of class 1 (PD). This suggests that although the model is highly accurate in predicting class 0, it faces challenges with class 1, likely due to class imbalance in the dataset.

Figure 26. The testing set confusion matrix for model-based features.



Discussion

This study presents a noninvasive approach for early detection of PD through the analysis of model-based gait features. Using kinematic characteristics, such as shoulder distance, step length, stride length, knee and hip angles, leg and arm symmetry, and trunk angles, we aimed to identify subtle gait abnormalities associated with PD. Data were collected through controlled video recordings of the TUG assessment, and the extracted features were processed using advanced filtering techniques and analyzed using SVM classifier.

The results demonstrate that the model-based features were highly effective in distinguishing between normal and PD-affected gait patterns, achieving high accuracy, precision, recall, and F_1-score. These findings support the potential of model-based gait analysis as a noninvasive, accessible tool for early diagnosis and monitoring of PD. The study also highlights the importance of addressing class imbalance and refining the feature extraction process to further enhance the performance of the classifier.

Future research should focus on validating these findings with larger datasets, exploring other machine learning models, and integrating additional features to improve the robustness and accuracy of PD detection systems. This approach could significantly benefit clinical settings, providing a reliable, noninvasive method for early diagnosis and improved patient outcomes.

Acknowledgments

This research is supported by the Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/MMU/02/5) and Malaysian-Jordan Matching Fund.

Conflicts of Interest

None declared.

References

1. Challa KN, Pagolu VS, Panda G, Majhi B. An improved approach for prediction of Parkinson's disease using machine learning techniques. In: Proceedings of the International Conference on Signal Processing, Communication, Power and Embedded System. 2016 Presented at: SCOPES 2016; October 3-5, 2016; Paralakhemundi, India. [doi: 10.1109/scopes.2016.7955679]
2. Polat K. Freezing of gait (FoG) detection using logistic regression in Parkinson's disease from acceleration signals. In: Proceedings of the Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science. 2019 Presented at: EBBT 2019; April 24-26, 2019; Istanbul, Turkey. [doi: 10.1109/ebbt.2019.8742042]
3. Vidya B, Sasikumar P. Gait based Parkinson’s disease diagnosis and severity rating using multi-class support vector machine. Appl Soft Comput 2021 Dec;113:107939. [doi: 10.1016/j.asoc.2021.107939]
4. Fang Z. Improved KNN algorithm with information entropy for the diagnosis of Parkinson's disease. In: Proceedings of the International Conference on Machine Learning and Knowledge Engineering. 2022 Presented at: MLKE 2022; February 25-27, 2022; Guilin, China. [doi: 10.1109/mlke55170.2022.00024]
5. Gundala S, Harichandana C, Jaasmitha A, Brundha PL. Parkinson’s disease identification by voice and handwritten drawings using Xgboost and Random Forest Algorithms. In: Proceedings of the 3rd International Conference on Computing and Information Technology. 2023 Presented at: ICCIT 2023; September 13-14, 2023; Tabuk, Saudi Arabia. [doi: 10.1109/iccit58132.2023.10273905]
6. Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Comput Sci 2023;218:249-261. [doi: 10.1016/j.procs.2023.01.007]
7. Pereira CR, Weber SA, Hook C, Rosa GH, Papa JP. Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: Proceedings of the 29th SIBGRAPI Conference on Graphics, Patterns and Images. 2016 Presented at: SIBGRAPI 2016; October 4-7, 2016; Sao Paulo, Brazil. [doi: 10.1109/sibgrapi.2016.054]
8. Grover S, Bhartia S, Akshama, Yadav A, KR S. Predicting severity of Parkinson’s disease using deep learning. Procedia Comput Sci 2018;132:1788-1794. [doi: 10.1016/j.procs.2018.05.154]
9. Aal HA, Taie S, El-Bendary N. An optimized RNN-LSTM approach for Parkinson’s disease early detection using speech features. Bull Electr Eng Inform 2021 Oct;10(5):2503-2512. [doi: 10.11591/eei.v10i5.3128]
10. Ouhmida A, Terrada O, Raihani A, Cherradi B, Hamida S. Voice-based deep learning medical diagnosis system for parkinson's disease prediction. In: Proceedings of the International Congress of Advanced Technology and Engineering. 2021 Presented at: ICOTEN 2021; July 4-5, 2021; Taiz, Yemen. [doi: 10.1109/icoten52080.2021.9493456]
11. Biswas S, Kaur N, KR S. Early detection of Parkinson’s disease from hand drawings using CNN and LSTM. In: Proceedings of the 4th International Conference on Artificial Intelligence and Speech Technology. 2022 Presented at: AIST 2022; December 9-10, 2022; Delhi, India. [doi: 10.1109/aist55798.2022.10065159]
12. Khaskhoussy R, Ayed YB. Improving Parkinson’s disease recognition through voice analysis using deep learning. Pattern Recognit Lett 2023 Apr;168:64-70. [doi: 10.1016/j.patrec.2023.03.011]

Abbreviations

ANN: artificial neural network
AUC: area under the curve
CNN: convolutional neural network

DNN: deep neural network
FoG: freezing of gait
KNN: k-nearest neighbors
LSTM: long short-term memory
PD: Parkinson disease
ROM: range of motion
SVM: support vector machine
TUG: timed up and go
UPDRS: Unified Parkinson Disease Rating Scale

Edited by G Seçkin; submitted 21.08.24; peer-reviewed by A Samraj, PS Teh, TT Verlekar; comments to author 18.11.24; revised version received 26.11.24; accepted 20.01.25; published 08.04.25.

Please cite as:
Lim MD, Connie T, Goh MKO, Saedon N'
Model-Based Feature Extraction and Classification for Parkinson Disease Screening Using Gait Analysis: Development and Validation Study
JMIR Aging 2025;8:e65629
URL: https://aging.jmir.org/2025/1/e65629
doi:10.2196/65629
PMID:

Original Paper

Identifying Deprescribing Opportunities With Large Language Models in Older Adults: Retrospective Cohort Study

Vimig Socrates[1,2*], MS; Donald S Wright[3,4*], MD, MHS; Thomas Huang[3], BS; Soraya Fereydooni[3], BS; Christine Dien[2], BS; Ling Chi[3], BS; Jesse Albano[5], PharmD, BCPS, BCEMP; Brian Patterson[6], MD, MPH; Naga Sasidhar Kanaparthy[3,4], MD, MPH; Catherine X Wright[7], MD; Andrew Loza[1,3], MD, PhD; David Chartash[1,8], PhD; Mark Iscoe[1,3], MD, MHS; Richard Andrew Taylor[1,3,9], MD, MHS

[1]Department of Biomedical Informatics and Data Science, School of Medicine, Yale University, New Haven, CT, United States
[2]Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT, United States
[3]Department of Emergency Medicine, School of Medicine, Yale University, New Haven, CT, United States
[4]VA Connecticut Healthcare System, US Department of Veterans Affairs, West Haven, CT, United States
[5]Department of Pharmacy, Yale New Haven Hospital, New Haven, CT, United States
[6]BerbeeWalsh Department of Emergency Medicine, University of Wisconsin–Madison, Madison, WI, United States
[7]Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, United States
[8]School of Medicine, University College Dublin, Dublin, Ireland
[9]Department of Biostatistics, School of Public Health, Yale University, New Haven, CT, United States
[*]these authors contributed equally

Corresponding Author:
Richard Andrew Taylor, MD, MHS
Department of Biomedical Informatics and Data Science
School of Medicine
Yale University
464 Congress Avenue
Suite 260
New Haven, CT, 06510
United States
Phone: 1 2037854058
Email: richard.taylor@yale.edu

Abstract

Background: Polypharmacy, the concurrent use of multiple medications, is prevalent among older adults and associated with increased risks for adverse drug events including falls. Deprescribing, the systematic process of discontinuing potentially inappropriate medications, aims to mitigate these risks. However, the practical application of deprescribing criteria in emergency settings remains limited due to time constraints and criteria complexity.

Objective: This study aims to evaluate the performance of a large language model (LLM)–based pipeline in identifying deprescribing opportunities for older emergency department (ED) patients with polypharmacy, using 3 different sets of criteria: Beers, Screening Tool of Older People's Prescriptions, and Geriatric Emergency Medication Safety Recommendations. The study further evaluates LLM confidence calibration and its ability to improve recommendation performance.

Methods: We conducted a retrospective cohort study of older adults presenting to an ED in a large academic medical center in the Northeast United States from January 2022 to March 2022. A random sample of 100 patients (712 total oral medications) was selected for detailed analysis. The LLM pipeline consisted of two steps: (1) filtering high-yield deprescribing criteria based on patients' medication lists, and (2) applying these criteria using both structured and unstructured patient data to recommend deprescribing. Model performance was assessed by comparing model recommendations to those of trained medical students, with discrepancies adjudicated by board-certified ED physicians. Selective prediction, a method that allows a model to abstain from low-confidence predictions to improve overall reliability, was applied to assess the model's confidence and decision-making thresholds.

Results: The LLM was significantly more effective in identifying deprescribing criteria (positive predictive value: 0.83; negative predictive value: 0.93; McNemar test for paired proportions: χ^2_1=5.985; *P*=.02) relative to medical students, but showed limitations in making specific deprescribing recommendations (positive predictive value=0.47; negative predictive value=0.93). Adjudication revealed that while the model excelled at identifying when there was a deprescribing criterion related to one of the patient's medications, it often struggled with determining whether that criterion applied to the specific case due to complex inclusion and exclusion criteria (54.5% of errors) and ambiguous clinical contexts (eg, missing information; 39.3% of errors). Selective prediction only marginally improved LLM performance due to poorly calibrated confidence estimates.

Conclusions: This study highlights the potential of LLMs to support deprescribing decisions in the ED by effectively filtering relevant criteria. However, challenges remain in applying these criteria to complex clinical scenarios, as the LLM demonstrated poor performance on more intricate decision-making tasks, with its reported confidence often failing to align with its actual success in these cases. The findings underscore the need for clearer deprescribing guidelines, improved LLM calibration for real-world use, and better integration of human–artificial intelligence workflows to balance artificial intelligence recommendations with clinician judgment.

(JMIR Aging 2025;8:e69504) doi:10.2196/69504

KEYWORDS

deprescribing; large language models; geriatrics; potentially inappropriate medication list; emergency medicine; natural language processing; calibration

Introduction

Polypharmacy, widely defined as the regular use of at least 5 medications, is common in older adults and at-risk populations [1]. In fact, approximately 30% of patients aged 65 years or older have polypharmacy [2], and nearly half of older emergency department (ED) patients are discharged with one or more new medications [3]. Although necessary and beneficial for some patients, polypharmacy can increase the risk of negative consequences for patients, including ED visits, adverse drug events, falls, disability, and inappropriate medication use [1]. While definitions differ, deprescribing is generally defined as a structured process by which potentially inappropriate medications (PIMs) are identified and withdrawn under the supervision of a health care provider. In some definitions, the process is described as evaluating the risk-benefit tradeoff, focusing on situations where the potential or actual harms of a medication outweigh its benefits, considering the patient's individual care goals and quality of life [2,4,5].

Deprescribing tools, such as the Screening Tool of Older People's Prescriptions (STOPP) [6] and Beers criteria [7], have been developed to help providers assess and identify PIMs based on a patient's medication list [7-9]. These explicit assessments are criterion-based with clear standards but are often impractical to implement in time-constrained clinical settings, such as ED, due to the need to evaluate multiple clinical indications and specialist-prescribed medications [10]. Attempts to digitize these criteria into electronic clinical decision support (CDS) have raised difficulties, typically requiring a labor-intensive coding process and unstructured information such as free text from patient records to contextualize certain criteria [11,12].

Large language models (LLMs) have been shown to interpret complex clinical situations and offer recommendations, from differential diagnoses to care management, leading to growing interest in their application in the medical field [13-16]. Moreover, they have been shown to extract medication-related data such as medication name, dosage, and frequency, necessary for the application of deprescribing criteria [17]. Finally, LLMs are excellent in-context learners, requiring very little labeled data to make predictions [18]. This reduces the annotation burden for time-constrained ED physicians while improving the use of unstructured patient records to contextualize patient medication lists. However, the majority of clinical reasoning evaluations on LLMs have been conducted using standardized exams (the United States Medical Licensing Examination) or digital case reports [14,19]. Their ability to perform clinical reasoning and calibrate responses over physician-generated text remains understudied.

Here, we propose to evaluate the performance of an LLM-based data pipeline in recommending deprescribing options for older adult ED patients at discharge based on 2 leading deprescribing criteria, Beers and STOPP. We have also included a recently developed list of criteria, Geriatric Emergency Medication Safety Recommendations (GEMS-Rx), intended to prevent the initiation of inappropriate medications in the acute care setting, as similar deprescribing lists specific to this care environment are not available [3]. Through this work, we hope to evaluate whether an LLM-based CDS system can effectively triage medications eligible for electronic deprescribing in older adults. Successful implementation of such a system would help address gaps in electronic deprescribing by using an LLM to contextualize recommendations within individual patient records and reduce manual development in CDS tools.

Methods

Ethical Considerations

This study was conducted with approval from the Institutional Review Board (IRB) at Yale University, under protocol number 2000035077. The IRB determined that this research qualifies for exemption as it involves secondary analysis of existing electronic health record (EHR) data, with no additional patient contact or data collection. The original data were collected with patient consent, and the current analysis adheres to the conditions of that consent and IRB approval, permitting

secondary analysis without the need for additional consent. All data were de-identified prior to analysis to ensure patient confidentiality. This study complies with ethical standards and guidelines for research involving human subjects.

Patient Cohort

All older adults (aged 65 years and older) with polypharmacy (5 or more active outpatient medications) presenting to an ED in a large academic medical center in the Northeast United States between January 2022 and March 2022 were identified. Due to budgetary constraints and a lack of prior evidence regarding the performance of LLMs in this task to guide a power calculation, we selected a random sample of 100 unique patients for evaluation.

Identification of Consensus-Based High-Yield Criteria

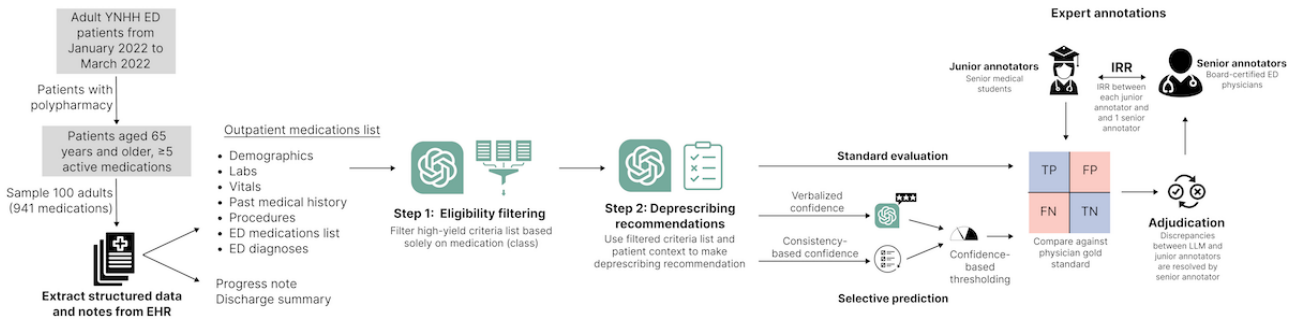
We conducted a consensus-based evaluation to filter three preexisting deprescribing lists (ie, STOPP [6], Beers [7], and GEMS-Rx [3]) into a focused set of high-yield deprescribing criteria for the LLM to use in its recommendations. High-yield criteria were defined as those posing a significant clinical risk to the patient and being identifiable within the electronic health records (EHRs). To identify these criteria, we evaluated 180 recommendations across 2 key dimensions: clinical risk and EHR computability. The consensus panel consisted of 6 board-certified physicians (in Emergency Medicine, Internal Medicine, and Clinical Informatics) and 1 ED pharmacist. Each member of the group individually reviewed each of the criteria and rated them on a 5-point Likert scale. We selected the top 50% of criteria with an average score greater than 3 on both dimensions, calculated across all experts, as high-yield criteria. We further elaborate on this consensus process and the final set of criteria in the Results section.

The need to filter the criteria before proceeding with the study was identified in our preliminary research [20], which revealed that one of the main causes of discrepancies between physicians and LLMs arose from ambiguous inclusion or exclusion conditions in deprescribing criteria. For example, criteria like "Statins for primary cardiovascular prevention in persons aged ≥85 with established frailty with expected life expectancy likely less than 3 years" include elements—such as "established frailty" and "expected life expectancy"—that are challenging to quantify and therefore difficult to implement computationally. The dimensions used to filter criteria were chosen to ensure that an LLM-enabled CDS tool prioritizes meaningful recommendations from high-quality EHR data, enabling accurate, actionable deprescribing recommendations and reduction of alert fatigue. We present the final set of high-yield criteria based on the average results of the consensus study.

Deprescribing Recommendations by GPT-4o

The study was approved under an exemption by the Yale University institutional review board prior to commencement (HIC# 2000035077). All patient-level data were deidentified prior to use with the LLM. We leveraged Microsoft's Azure OpenAI GPT-4o (GPT-4o model version: 2024-08-06 and OpenAI API version: 2024-02-15-preview) to produce deprescribing recommendations through a 2-stage process, as shown in Figure 1. In stage 1, GPT-4o was prompted to filter the full list of high-yield criteria solely based on the patient's medication list, ignoring inclusion or exclusion conditions. In stage 2, GPT-4o was prompted to use its previously filtered criteria list, along with structured (eg, demographics, lab values, vitals, and past medical history) and unstructured (most recent progress note and discharge summary) information, to determine if the patient satisfied any deprescribing criteria and the medication should be recommended for deprescribing. Each medication was evaluated individually to prevent errors from simultaneous processing, such as misattribution of criteria or medication omissions. To ensure optimal performance, we engineered prompts in an iterative fashion [21], using 1 patient at a time from a set of patients (up to 10% of the cohort) not used in the subsequent evaluation. After each evaluation, prompts were adjusted to correct any systematic errors (eg, instances where no relevant criteria in step 1 led to noncriteria-based deprescribing recommendations in step 2) by the LLM. After our third prompt yielded an output without any identifiable errors, we stopped the iterative prompt development process. Consequently, the final 2 patients initially reserved for this purpose were included in the final cohort evaluation (n=92 patients, 626 medications). Aside from the consistency-based method described later, all LLM calls were performed with a fixed temperature (temperature=0; low randomness in generated responses) and seed to ensure reproducibility and deterministic outputs.

Figure 1. Overview of the evaluation pipeline, consisting of a 2-step GPT-4o process, performance comparison with junior annotators (medical students), and final adjudication by senior annotators (board-certified physicians). ED: emergency department; EHR: electronic health record; FN: false negative; FP: false positive; IRR: interrater reliability; TN: true negative; TP: true positive; YNHH: Yale New Haven Hospital.



This 2-stage process was developed to correct errors identified when both stages were accomplished at once. Our initial testing revealed that providing the LLM with the full set of criteria and medications led to simple reasoning errors. The large number

of criteria, combined with the simultaneous processing of the complete patient medication list, resulted in inaccuracies in applying individual criteria to specific medications. Separating the process of criteria filtering and application both reduces confusion due to large input context sizes [22] and ensures that extraneous context does not distract the LLM [23]. The full prompts for step 1 and step 2 are included in Figures S5 and S6 in Multimedia Appendix 1.

Selective Prediction Methods

In addition to evaluating an LLM's ability to make deprescribing recommendations, we assessed whether its confidence estimates were well-calibrated and examined their impact on predictive performance. To do so, we collected GPT-4o's decision confidence for both steps using 2 validated confidence elicitation methods: chain-of-thought verbalized confidence and self-random sampling with average-confidence aggregation, referred to as consistency-based confidence [24]. In verbalized confidence, we asked the LLM to explicitly estimate its confidence for each step following its decision. For the consistency-based approach, we followed the best practices established in prior work [24] by sampling the LLM multiple times (number of samples=5) with high temperature (T=0.8; high randomness in generated responses) and used a majority vote weighted by the confidence of each response to determine the final deprescribing recommendation.

In a human-in-the-loop decision-making system, the LLM's confidence would be used to determine if the model should abstain due to low certainty regarding its own decision. In practice, this case would be considered too difficult for the LLM and forwarded to an expert reviewer. This human-in-the-loop decision-making pipeline is known as selective prediction and has been commonly found to improve performance in non–text-based applications [25,26]. We evaluated both selective prediction methods using risk-coverage curves [27,28], substituting risk for the F_1-score (a measure of the predictive performance of a model balancing precision and recall) to capture the full range of predictive performance. Coverage was also expressed inversely as the deferring fraction, representing the proportion of instances where the LLM abstained from making a decision. We conducted a more in-depth analysis of the method that proved to be more effective.

Comparison and Adjudication With Clinical Experts

In this study, we used a rigorous human review and adjudication process to assess model performance. Two trained senior medical students (M4) classified all medications in the test cohort using a 2-stage pipeline, after first computing interrater reliability (IRR) on an adjudication set of 75 medications from 5 patients. Similar to the LLM pipeline, for each medication, a medical student determined (1) if there exists a relevant high-yield criteria based on the medication list, and (2) whether the medication should be recommended for deprescribing. Discrepancies between the students (junior annotators) and the LLM across both stages were adjudicated by 2 board-certified ED physicians (senior annotators). Similarly to the junior annotators, we measured the IRR between the 2 senior ED physicians, prior to adjudication on the full set of discrepancies. Finally, we classified the errors leading to incorrect recommendations by the LLM, leveraging a prior evaluation framework [29]. We classified each error as 1 of 4 error types: incorrect reading comprehension, incorrect recall of knowledge, incorrect reasoning step, and not enough information, as described in Table 1.

Table 1. Definitions of GPT-4o error types inspired by framework from Lièven et al [29] relevant to deprescribing recommendations.

GPT-4o error types	Definition	Example
Incorrect reading comprehension	Includes misunderstanding of order of text, such as when a medication is dependent on another medication in a specific arrangement. Also includes ignoring information provided in the input text, such as missing a relevant category explicitly stated in the recommendations.	GPT failed to recognize acetaminophen by name from the list of STOPP[a] criteria in a patient at risk for malnutrition or liver disease.
Incorrect recall of knowledge	Includes failure to recognize classes of medications or other medical facts necessary to perform the task.	GPT correctly recognized amlodipine was a calcium channel blocker but failed to recognize it was more broadly an antihypertensive.
Incorrect reasoning step	Faulty reasoning, such as inappropriate assumptions or leads of logic unsupported by the clinical data.	GPT recommended discontinuing warfarin in a patient with a therapeutically elevated INR[b] after assuming that this elevated INR was due to a bleeding disorder.
Not enough information	Inappropriate application of missing data leading to potentially unreliable conclusions, such as assuming abnormality of a missing laboratory study.	GPT recommended discontinuing a QT[c] prolonging antidepressant based on the possibility of QT prolongation without any ECG[d] data or history of abnormal QT interval.

[a]STOPP: Screening Tool of Older People's Prescriptions.

[b]INR: International normalized ratio.

[c]QT: QT interval.

[d]ECG: electrocardiogram.

Data Analysis

We evaluated whether the LLM or the medical student was correct, using a gold standard derived from senior annotator (board-certified ED physicians) adjudication of discrepancies. To compare their proportions of correct responses, we applied the McNemar test, a statistical method commonly used to analyze paired nominal data, such as diagnostic accuracy from different assessments applied to the same cases [30]. All analysis was performed using Python (version 3.9; Python Software Foundation). Statistical testing was carried out using *statsmodels* (version 0.14.4) [31] and all visualizations were generated using *seaborn* (version 0.13.2) [32] and *matplotlib* (version 3.8.2) [33].

Results

Patient Cohort

In total, we identified 10,977 unique patients across 15,161 emergency department encounters from January 2022 to March 2022 meeting our selection criteria, from which 100 patients were randomly selected (Table 2). As our criteria only pertain to oral medication, nonoral medications were subsequently filtered out, resulting in 712 total oral medications across the cohort and a median of 6 oral medications per patient (Figure 2). From our initial study cohort of 100 patients, 10 patients were set aside for both prompt engineering and calculation of IRR between junior annotators. Fewer iterations were needed to refine the prompt than initially anticipated, so the remaining 2 patients were included in the final study cohort. This resulted in a final evaluation cohort of 92 patients, encompassing a total of 626 medications. Based on the mechanism of action, statins were the most common medication class (atorvastatin and rosuvastatin; 6.8% combined), followed by proton pump inhibitors (pantoprazole, esomeprazole, and omeprazole; 4.6% combined). When classified by therapeutic effect, antihypertensive agents were most prevalent (amlodipine, lisinopril, losartan, and valsartan; 10.6% combined).

Table 2. Demographic overview of the 100 patients included in the evaluation (N=100).

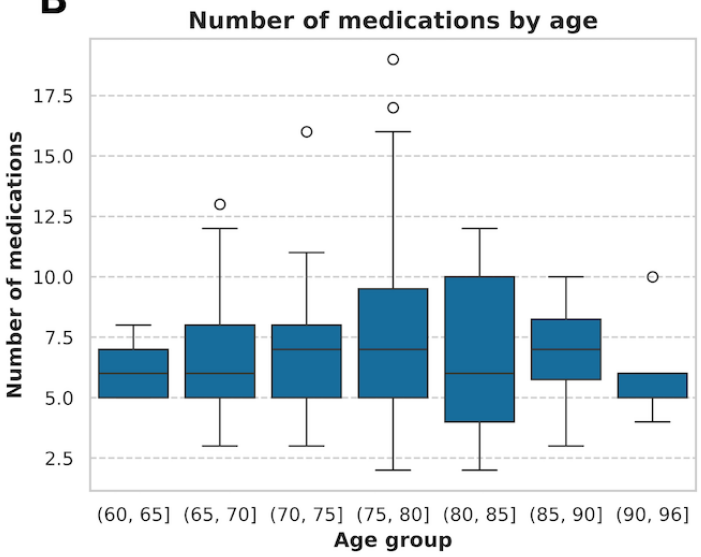
Characteristics	Values
Age, mean (SD)	75.8 (7.6)
Sex, n (%)	
Female	63 (63)
Male	37 (37)
Race, n (%)	
Asian	1 (1)
Black or African American	16 (16)
White or Caucasian	69 (69)
Other or not listed	13 (13)
None	1 (1)
Ethnicity, n (%)	
Hispanic or Latino	10 (10)
Non-Hispanic	90 (90)
Smoking status, n (%)	
Former smoker	47 (47)
Never smoker	43 (43)
Current every day smoker	8 (8)
Passive smoke exposure—never smoker	1 (1)
Light tobacco smoker	1 (1)
Number of medications, median (IQR)	6.0 (5.0-8.2)

Figure 2. Medication information about baseline cohort. (A) The top 20 most commonly prescribed medications represented as a percentage of the total medication set. (B) Distribution of medications across different age groups.

A

Medication name	Total medications (%)
Amlodipine	4.6
Atorvastatin	4.2
Metoprolol succinate	3.3
Aspirin	3.3
Lisinopril	2.6
Rosuvastatin	2.6
Hydrochlorothiazide	2.4
Losartan	2.1
Gabapentin	2.1
Famotidine	1.8
Prednisone	1.8
Escitalopram	1.8
Sennosides	1.6
Lorazepam	1.6
Pantoprazole	1.6
Esomeprazole	1.5
Apixaban	1.5
Acetaminophen	1.5
Omeprazole	1.5
Valsartan	1.3

B

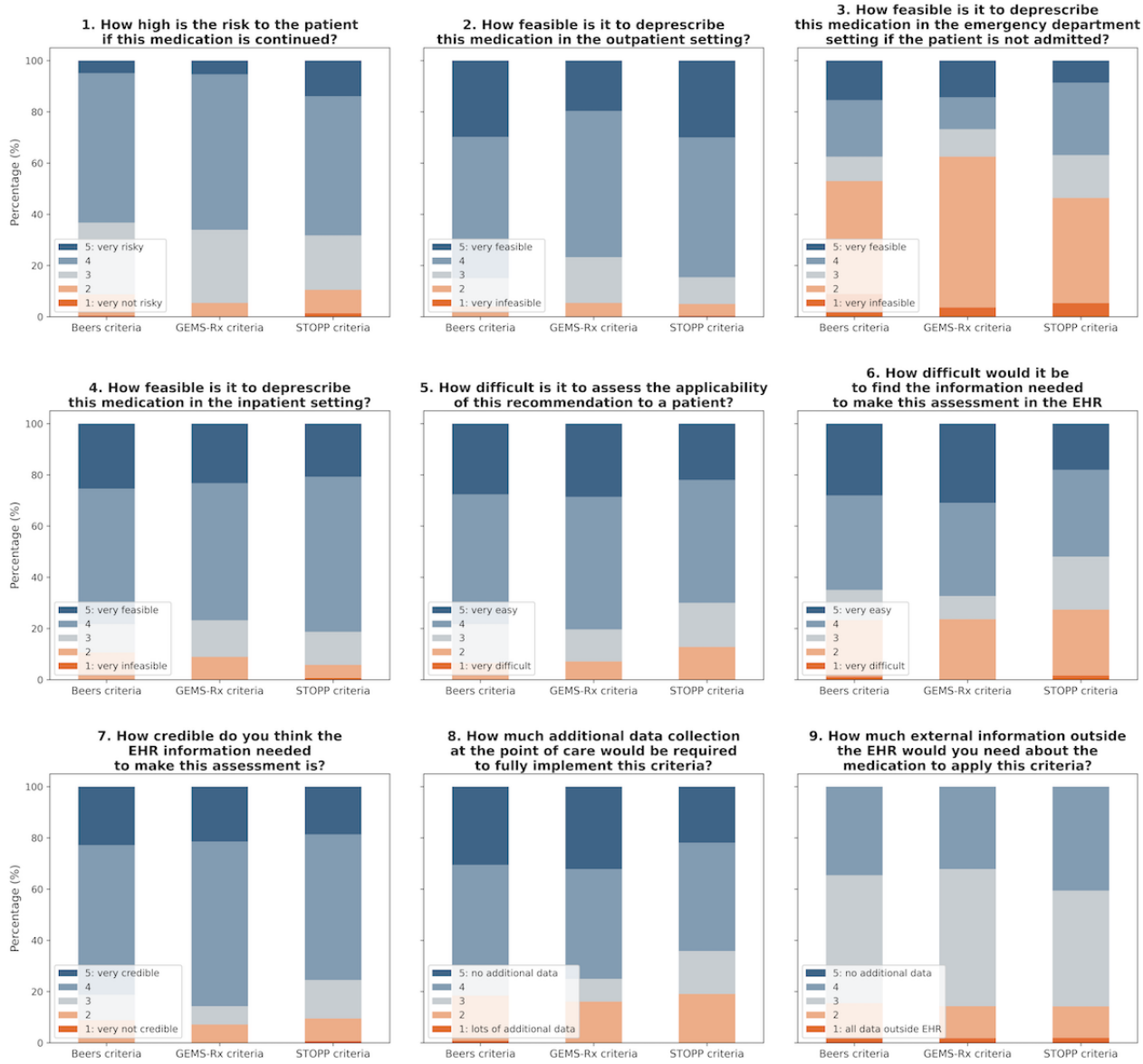


Evaluation of Consensus-Based High-Yield Criteria

To streamline evaluation by the LLM, we filtered criteria using the average scores from an expert consensus panel. All criteria (n=180) were evaluated based on their scores for clinical risk (Q1) and EHR computability (Q2-Q5), as assessed by an expert panel; average scores by deprescribing list are shown in Figure 3. Any criteria scoring less than 3 on both clinical applicability and EHR computability were excluded resulting in 161 criteria. From the remaining set, the top 50% were selected, resulting in 81 high-yield criteria across all 3 deprescribing lists.

On average, STOPP criteria had the lowest clinical risk and EHR computability ratings, while the Beers criteria had the highest, contributing to their respective adoption rates of 45.9% and 62.5% among high-yield criteria. Reduced inclusion of STOPP criteria was primarily attributed to panelists' concerns that the information required was not readily accessible within the EHR and would necessitate additional data at the point of care. We also present the results of feasibility in various clinical settings by criteria list in Figure S1 in Multimedia Appendix 1, showing more likelihood for deprescribing in both outpatient and inpatient contexts, compared to the ED, across all 3 criteria lists. A scatter plot of risk versus EHR computability of high-yield criteria is presented in Figure S2 in Multimedia Appendix 1.

Figure 3. Average distribution of results used to filter high-yield criteria on a 5-point Likert scale from a consensus study by an expert panel (n=7) split by 3 criteria lists: Beers, GEMS-Rx, and STOPP. EHR: electronic health record; GEMS-Rx: Geriatric Emergency Medication Safety Recommendations; STOPP: Screening Tool of Older People's Prescriptions.



Deprescribing Recommendations by GPT-4o

We next evaluated the LLM's deprescribing recommendations by comparing them to those of medical students, resolving any discrepancies through adjudication by board-certified EM physicians. As shown in Figure 4, 315 medications (50.3% of the total) lacked relevant high-yield deprescribing criteria. The LLM effectively identified these cases, achieving an F_1-score of 0.86 (precision=0.83, recall=0.90). Among those medications with relevant criteria, 64 (10.2% of the total) were cases where either GPT-4o or the medical student recommended deprescribing. In this second step, which involved applying relevant criteria to make a recommendation, the LLM performed less effectively, with an F_1-score of 0.58 (precision=0.47, recall=0.76).

For cases where the LLM and the medical students differed, 2 senior annotators (board-certified Emergency Medicine physicians) adjudicated 126 discrepancies after standardizing the codebook and verifying IRR (Cohen *k*: eligibility=0.795, deprescribing=0.745). Notably, the confusion matrix (Figure 4) revealed that a major source of discrepancy was the significantly higher likelihood of the LLM to recommend deprescribing (11.6%) compared to the medical students (1.91%). The confusion matrix describes all possible outcomes when comparing the LLM with the medical students (eg, the medical student recommended deprescribing with eligible criteria and the LLM found no eligible criteria). The adjudication yielded similar results to those observed in comparison with medical students (Figure 5A). Across all discrepancies, GPT-4o was significantly more effective in criteria filtering compared to the medical student (McNemar test for paired proportions [30]: χ^2_1=5.985; *P*=.015). However, in applying relevant criteria, GPT-4o performed worse than the medical students, though the difference was not statistically significant (McNemar test: χ^2_1=1.818; *P*=.178). Adjudication was chosen as the gold standard for resolving discrepancies because preliminary

research showed low IRR among medical students for both steps of the process (Cohen *k*: step 1=0.68, step 2=0.33) across 75 medications.

In cases where GPT-4o was incorrect, error analysis highlighted key patterns across the 2 steps. For criteria filtering, senior annotators observed that errors often stemmed from incorrect reasoning or reading comprehension issues (Figure 5B). Similarly, in making recommendations, the most common error was incorrect reasoning, followed by cases where the LLM lacked sufficient information to make an accurate determination and subsequently made inappropriate assumptions.

Figure 4. Confusion matrix (n=626 medications). The joint confusion matrix across both steps showing alignment and discrepancies between the GPT-4o model and medical students.

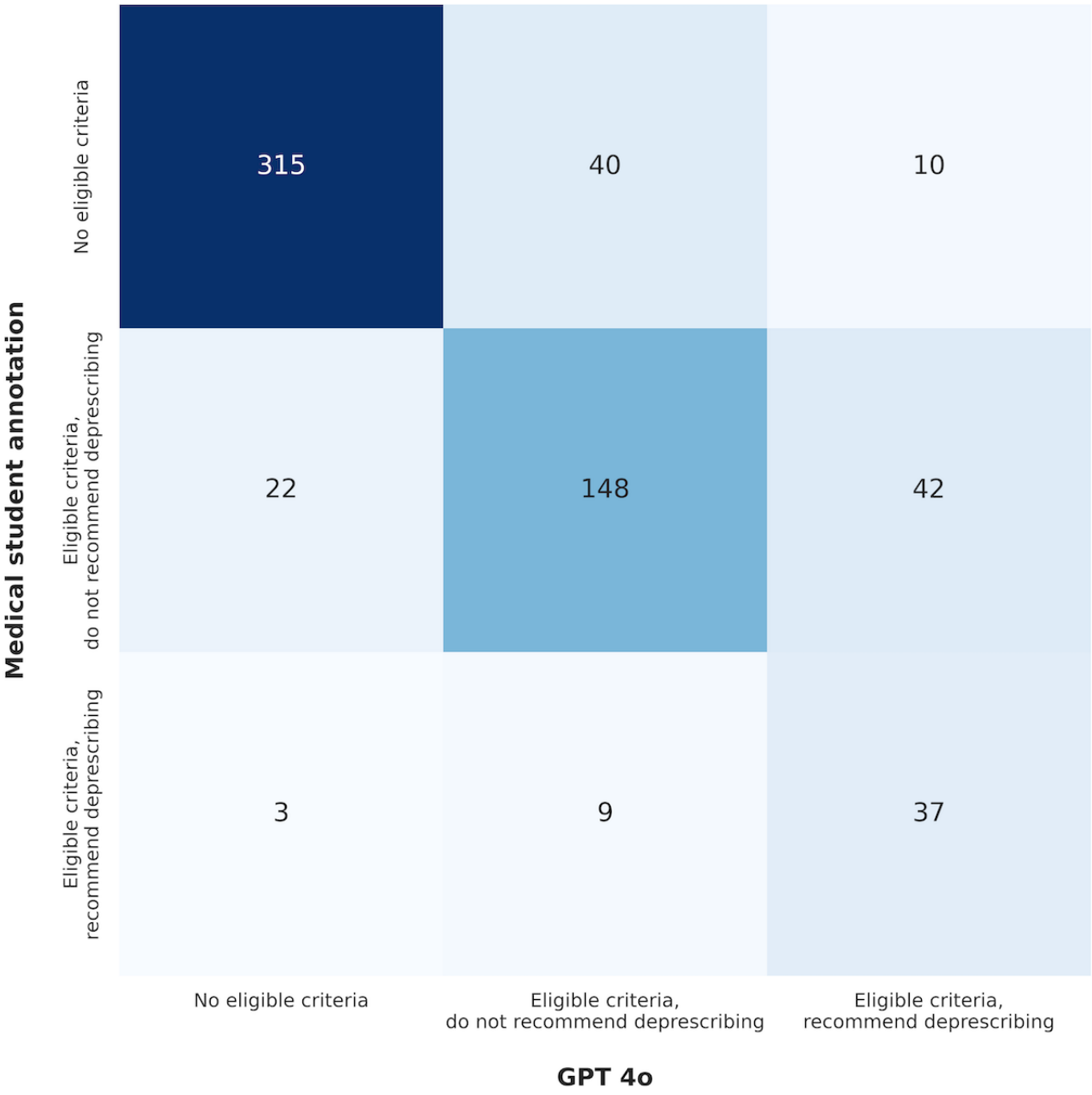
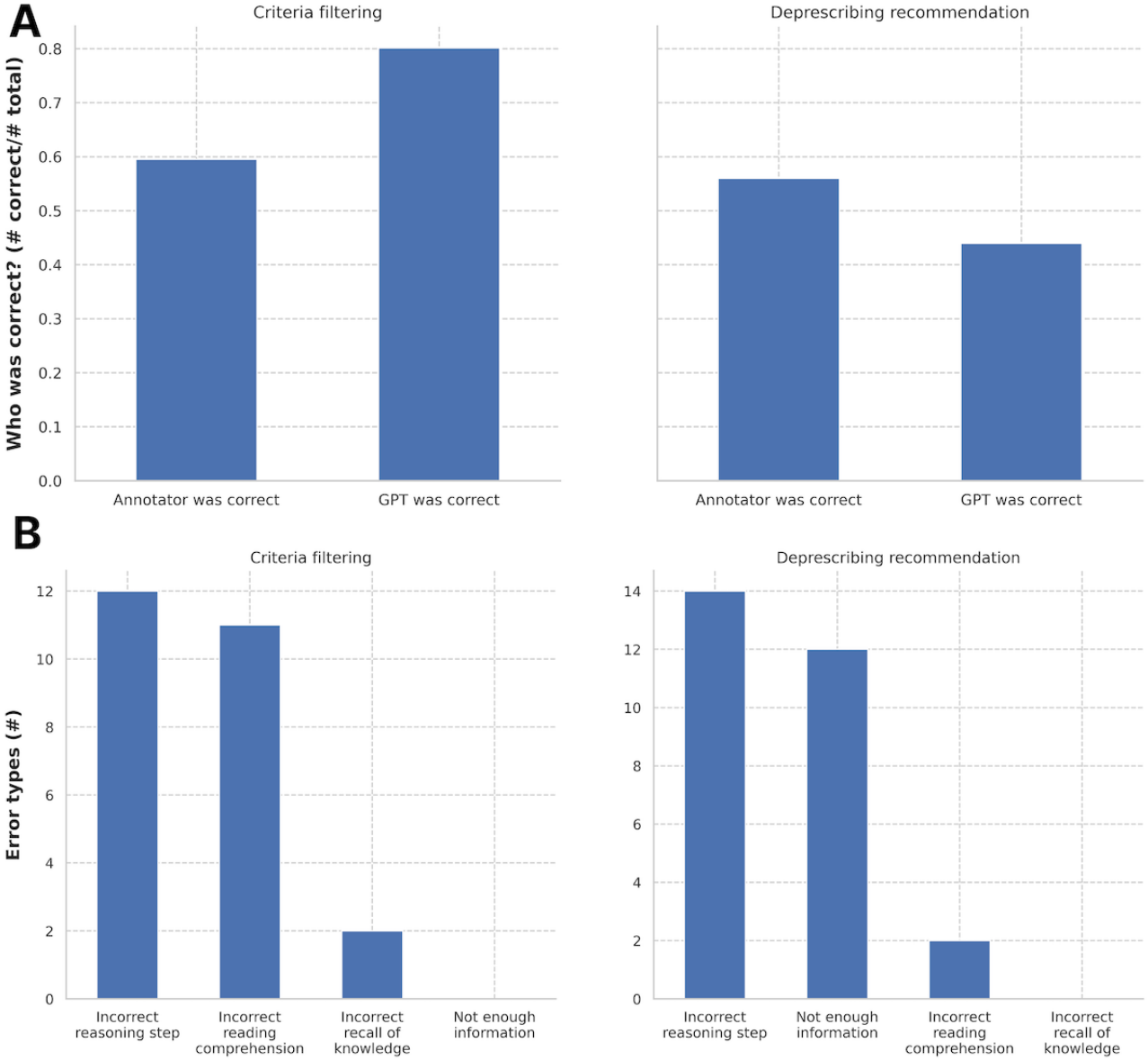


Figure 5. (A) Expert adjudication (n=126). Adjudication by senior clinical expert comparing junior annotators and GPT-4o in both criteria filtering and deprescribing recommendation tasks. (B) GPT-4o error modes. Types of errors by GPT-4o in the adjudication set (n=126) for both criteria filtering and deprescribing recommendation tasks.

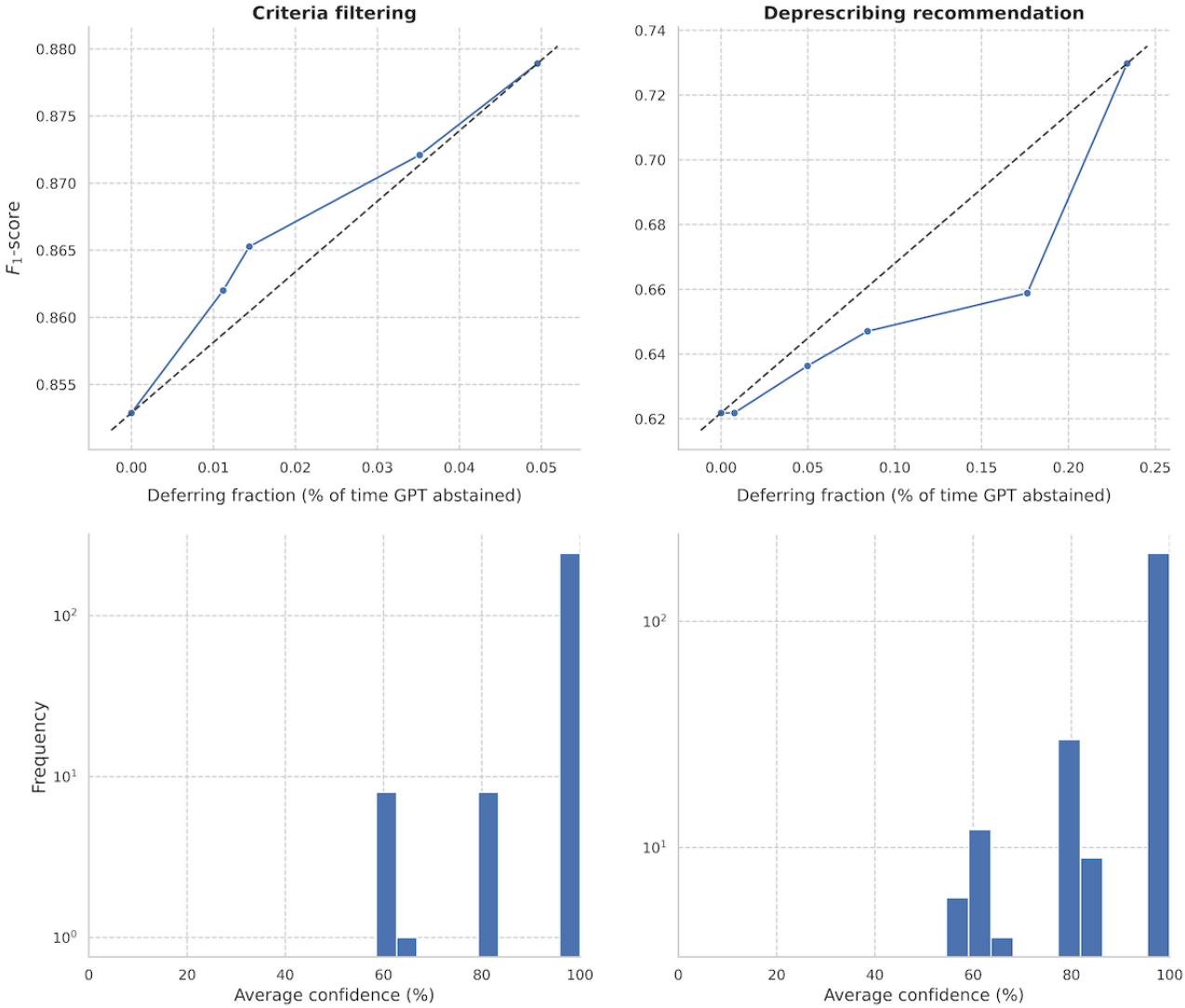


Selective Prediction Methods

Finally, we investigated the impact of incorporating confidence estimates from the LLM to guide selective prediction, allowing the model to abstain from making predictions in cases of low confidence. We compared 2 approaches: verbalized confidence and consistency-based confidence. Verbalized confidence demonstrated a narrower range of F_1-scores overall, with step 1 (eligibility) ranging from 0.860 to 0.863 and step 2 (deprescribing) from 0.58 to 0.69, as shown in Figure S3 in Multimedia Appendix 1. In contrast, consistency-based confidence exhibited broader and higher F_1-score ranges, with step 1 spanning 0.85 to 0.88 and step 2 ranging from 0.62 to 0.73 (Figure 6A). These results suggest that consistency-based confidence provides more flexibility and improved performance compared to verbalized confidence across both steps.

In particular, consistency-based selective prediction demonstrates a positive linear relationship between accuracy in deprescribing recommendations and deferring fractions. However, despite some minor improvements, we find that the LLM is poorly calibrated, as shown in Figure 6B. Despite consistency-based weighting, the confidence distribution is severely left-skewed with the minimum confidence being 58.5% in eligibility filtering and 54.5% in deprescribing recommendations.

Figure 6. Consistency-based selective prediction. (A) Range of F1-scores resulting from the application of consistency-based selective prediction in both steps of the deprescribing pipeline. The dotted line shows ideal performance as a function of deferring fraction. (B) Distribution of confidences for deprescribing recommendations from GPT-4o on a log scale.



Discussion

Principal Findings

In this retrospective cohort study evaluating deprescribing opportunities for PIMs among older adults with polypharmacy in the ED, we found that LLMs effectively identify relevant criteria from verified lists but are less adept at applying these criteria to individual patient cases. GPT-4o's performance was compared to that of medical students in a 2-step pipeline: filtering for criteria-eligible medications and making specific deprescribing recommendations. Adjudication by senior clinicians was used to resolve discrepancies, and selective prediction methods were tested to improve the model's reliability. The results offer insights into both the capabilities and limitations of LLMs in a real-world clinical context, highlighting key areas for improvement in both LLM frameworks and deprescribing guidelines.

Effectiveness of the 2-Step LLM Pipeline

The LLM demonstrated strengths in the initial filtering step, accurately identifying a high proportion of medications that matched deprescribing criteria, thus offering the potential to support clinicians in rapidly screening complex medication lists. In fact, the LLM outperformed medical students by a significant margin (80.1% vs 59.5% correct, McNemar test: P=.02). The adjudication, combined with strong overall performance (maximum F_1-score: 87.8%) using selective prediction methods, suggests that the LLM can effectively minimize the number of criteria requiring final review for deprescribing recommendations. Cases of misclassification were relatively uncommon and primarily related to nonstandard drug class names or overly broad groups, which could be improved with refined deprescribing criteria. However, in the second step—making specific deprescribing recommendations—the LLM encountered considerable difficulty, particularly when dealing with ambiguous criteria, missing information, and nuanced clinical scenarios. For example, thiazide diuretics are recommended for deprescribing in cases of significant hypokalemia, hyponatremia, hypercalcemia, or a history of gout. However, GPT-4o recommended deprescribing without access to current electrolyte values, instead basing its suggestion on a history of chronic kidney disease, a condition associated with

potential electrolyte imbalances but not meeting the relevant inclusion criteria. If implemented in clinical decision support, these inaccuracies might contribute to increased alert fatigue and extend the time required to interpret LLM-generated recommendations, potentially offsetting the intended efficiency gains in identifying deprescribing opportunities.

Our findings on the LLM's performance in identifying medications with relevant deprescribing criteria based on eligibility guidelines align with evidence from clinical trial matching literature [34], where LLMs have shown performance comparable to physicians in applying such criteria to identify eligible patients. One study used a similar "filter-and-apply" pipeline, in which trials were first filtered and then matched to patients, showcasing the effectiveness of this approach [35]. However, despite successes in eligibility filtering, challenges remain when applying complex criteria to specific patient cases. Similar errors to those observed in our work have been reported, such as incorrectly identifying patients who meet partial criteria, or assuming a patient with breast cancer does not have lung cancer simply because it is not explicitly mentioned [36,37]. Overall, while LLMs hold promise for reducing the time burden in determining deprescribing eligibility, their application requires careful consideration, particularly in tasks involving complex clinical reasoning.

Role of Selective Prediction in Clinical Decision-Making

To address the model's limitations in clinical decision-making, we implemented selective prediction methods, which allowed the LLM to "abstain" from making a recommendation in cases of low confidence. Selective prediction marginally improved the LLM's filtering accuracy by enabling it to defer uncertain cases to human reviewers. However, the effectiveness of this approach was limited by the poorly calibrated confidence levels assigned by the LLM to its decisions. Specifically, the LLM displayed a minimum confidence level of 54%, even in cases where its recommendations were incorrect. This indicates a tendency toward overconfidence, particularly in its deprescribing recommendations. While verbalized confidence is known to be overconfident in clinical question answering [38], our results contradict recent work that suggests that consistency-based methods alleviate some of these concerns [39]. This discrepancy underscores the importance of task-specific confidence thresholds and suggests that selective prediction, while useful, is not a one-size-fits-all solution in complex clinical applications.

Improved uncertainty calibration in LLMs could enhance selective prediction methods, optimizing physician-artificial intelligence (AI) workflows in clinical settings. Future applications of a well-calibrated deprescribing CDS tool could flag cases where critical information is missing (eg, antipsychotics at unchanged doses for more than 3 months without a documented medication review). This approach could streamline medication filtering while preserving human oversight, allowing clinicians to focus on complex cases where LLM reliability is limited.

Need for Clearer Deprescribing Guidelines

A notable finding from this study is the need for clearer and more consistent deprescribing guidelines. Ambiguities in criteria definitions, such as those related to medication administration routes and drug classes, present substantial barriers to automation and contribute to discrepancies between human and model interpretations. Additionally, the model often recommended deprescribing medications that, while potentially inappropriate, required specific contextual qualifiers (eg, patient's life expectancy, nutritional status, frailty status) to justify deprescribing—criteria that the LLM misapplied due to lack of explicit context or ambiguous language in the guidelines. It is important to note that current deprescribing criteria were not originally designed for direct implementation in CDS systems, but rather as general recommendations for prescribing physicians. Reorganizing these criteria into a structured, explicit framework tailored for CDS use could reduce ambiguity, improve the model's performance, and support more consistent application in clinical practice. In general, streamlining deprescribing criteria to ensure consistent application across clinical contexts could improve model reliability and help standardize deprescribing practices.

Implications for LLM Use in Clinical Practice and Future Directions

The results of this study underscore the promise of LLMs in enhancing deprescribing workflows by providing rapid filtering of PIMs, which could alleviate some of the burdens on health care providers. However, this work also highlights the limitations of current LLMs in complex, context-sensitive clinical decision-making tasks. The LLM's frequent tendency to overrecommend deprescribing, as compared to medical students, indicates that clear boundaries for medication eligibility and exclusion are critical for reducing false positives in automated recommendations. Cases with the potential for human harm were observed, such as suggesting deprescribing anticoagulation in a patient with recent thromboembolism. Additionally, cases were seen in which the LLM recommended deprescribing without citing a specific criterion. These behaviors suggest that strong guardrails on the LLM are needed to ensure safe, high-quality recommendations. Enhancing guideline specificity, particularly for complex inclusion or exclusion criteria, could reduce both human and model error rates and may foster greater acceptance of AI-assisted deprescribing tools among clinicians. Our findings highlight the potential value of human-AI collaboration frameworks. For example, a human-in-the-loop framework (a model in which humans review difficult cases the LLM cannot resolve) could involve LLMs assisting in the identification of deprescribing opportunities while deferring final recommendations to clinicians [40]. Alternatively, the LLM could focus on the initial filtering of relevant deprescribing criteria for specific medications, leaving the recommendation task entirely to the clinician, thereby leveraging the LLM's strength in mapping medications or medication classes to appropriate criteria efficiently [41]. These approaches not only leverage the model's efficiency in data processing but also mitigate risks associated with erroneous recommendations, particularly in this high-risk clinical context. Future research should focus on refining LLM architectures to better handle the nuances of clinical reasoning and context interpretation, perhaps by incorporating more advanced natural language processing techniques and domain-specific training.

Additionally, efforts to standardize deprescribing guidelines would greatly benefit the development of automated tools in this area, making them more reliable and broadly applicable.

Limitations

This study has several limitations that warrant consideration. First, the retrospective nature of our analysis, relying on historical data from EHRs, may not fully capture the complexity of real-time clinical decision-making in emergency settings. The study's focus on a single large academic medical center limits the generalizability of our findings to other settings with different patient populations, documentation patterns, and health care practices. Second, the selective prediction methods, while providing insights into the LLM's confidence, were not universally effective, particularly in the nuanced task of deprescribing recommendations. The model's performance in these recommendations highlights the challenge of translating structured criteria into actionable clinical decisions, especially when faced with ambiguous inclusion or exclusion conditions. Additionally, the model's reliance on textual prompts and structured EHR data may not fully account for nuanced clinical contexts that influence deprescribing decisions. Third, the small sample size for detailed analysis (100 patients) limits the statistical power and may not reflect broader patterns of medication use and deprescribing needs. Cost may be a barrier to larger sample sizes in the future, as the total application programming interface utilization fees were approximately US $300-$400 over these 100 patients and the cost to both evaluate and implement the system in the real world would scale linearly with the study population. Additionally, the study relied on medical students for initial annotation. While these annotations were reviewed and adjudicated by board-certified physicians, this process may introduce variability, potentially affecting the reliability of their use as a gold standard in selective prediction methods. Our process for selecting these criteria also did not explicitly include any geriatricians, though did include a range of individuals who regularly care for older adults. Finally, the criteria used (STOPP, Beers, and GEMS-Rx) were selected based on their perceived clinical risk and EHR computability, which may not encompass all relevant deprescribing scenarios. The lack of standardized guidelines for implementing deprescribing criteria in LLMs also poses a challenge to consistency and accuracy.

Conclusions

This study demonstrates the potential of LLMs to augment clinical decision support by effectively filtering deprescribing criteria for older adults with polypharmacy in ED. While the LLM showed promise in identifying medications eligible for deprescribing, it faced challenges in making nuanced deprescribing recommendations, underscoring the need for human oversight in AI-driven processes. Future research should prioritize refining the model by addressing ambiguities in deprescribing criteria and integrating broader clinical context, such as longitudinal data from prior progress notes and discharge summaries, to enable the detection of relevant clinical trends. Expanding the dataset and exploring more effective strategies for integrating human judgment with AI capabilities will help overcome limitations in generalizability, helping optimize patient care. The findings underscore the potential of LLMs in AI-enabled automated CDS tools for deprescribing while emphasizing the need to refine deprescribing criteria and establish clearer guidelines to support the integration of AI into clinical practice.

Acknowledgments

VS has received funding from the Yale University-Mayo Clinic Center of Excellence in Regulatory Science and Innovation for this work. MI receives funding through CTSA (grant UL1 TR001863) from the National Center for Advancing Translational Science, a component of the National Institutes of Health. This publication's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

VS, MI, and RAT conceived of the study with help from DSW in the design of its analysis. VS, CD, and LC provided data engineering and analyzed the data. VS prepared figures. DSW, JA, BP, NSK, CXW, AL, and RAT participated in the consensus study. TH, SF, DSW, and MI annotated notes for model training and testing. VS, DSW, LC, and RAT drafted the manuscript, and all authors contributed substantially to its revision. RAT takes responsibility for the paper as a whole.

Conflicts of Interest

RAT receives unrelated support from grants from the National Institutes of Health, Gordon and Betty Moore Foundation, the Federal Drug Administration, the Agency for Healthcare Research & Quality, and Beckman Coulter, Inc as well as options from Vera Health for serving as an advisor. DC is the associate editor for *JMIR Medical Education*.

Multimedia Appendix 1
Additional figures and tables.
[DOCX File , 1455 KB - aging_v8i1e69504_app1.docx]

References

1. Halli-Tierney AD, Scarbrough C, Carroll D. Polypharmacy: evaluating risks and deprescribing. Am Fam Physician 2019;100(1):32-38 [FREE Full text] [Medline: 31259501]
2. Scott IA, Hilmer SN, Reeve E, Potter K, Le Couteur D, Rigby D, et al. Reducing inappropriate polypharmacy: the process of deprescribing. JAMA Intern Med 2015;175(5):827-834. [doi: 10.1001/jamainternmed.2015.0324] [Medline: 25798731]
3. Skains RM, Koehl JL, Aldeen A, Carpenter CR, Gettel CJ, Goldberg EM, et al. Geriatric emergency medication safety recommendations (GEMS-Rx): modified Delphi development of a high-risk prescription list for older emergency department patients. Ann Emerg Med 2024;84(3):274-284. [doi: 10.1016/j.annemergmed.2024.01.033] [Medline: 38483427]
4. Reeve E, Gnjidic D, Long J, Hilmer S. A systematic review of the emerging definition of 'deprescribing' with network analysis: implications for future research and clinical practice. Br J Clin Pharmacol 2015;80(6):1254-1268 [FREE Full text] [doi: 10.1111/bcp.12732] [Medline: 27006985]
5. Hohl CM, Dankoff J, Colacone A, Afilalo M. Polypharmacy, adverse drug-related events, and potential adverse drug interactions in elderly patients presenting to an emergency department. Ann Emerg Med 2001;38(6):666-671. [doi: 10.1067/mem.2001.119456] [Medline: 11719747]
6. O'Mahony D, Cherubini A, Guiteras AR, Denkinger M, Beuscart JB, Onder G, et al. STOPP/START criteria for potentially inappropriate prescribing in older people: version 3. Eur Geriatr Med 2023;14(4):625-632 [FREE Full text] [doi: 10.1007/s41999-023-00777-y] [Medline: 37256475]
7. American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2023 updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. J Am Geriatr Soc 2023;71(7):2052-2081. [doi: 10.1111/jgs.18372] [Medline: 37139824]
8. Candeias C, Gama J, Rodrigues M, Falcão A, Alves G. Potentially inappropriate medications and potential prescribing omissions in elderly patients receiving post-acute and long-term care: application of screening tool of older people's prescriptions/screening tool to alert to right treatment criteria. Front Pharmacol 2021;12:747523 [FREE Full text] [doi: 10.3389/fphar.2021.747523] [Medline: 34737705]
9. Kaufmann CP, Tremp R, Hersberger KE, Lampert ML. Inappropriate prescribing: a systematic overview of published assessment tools. Eur J Clin Pharmacol 2014;70(1):1-11. [doi: 10.1007/s00228-013-1575-8] [Medline: 24019054]
10. Lee S, Bobb Swanson M, Fillman A, Carnahan RM, Seaman AT, Reisinger HS. Challenges and opportunities in creating a deprescribing program in the emergency department: a qualitative study. J Am Geriatr Soc 2023;71(1):62-76 [FREE Full text] [doi: 10.1111/jgs.18047] [Medline: 36258309]
11. Scott IA, Pillans PI, Barras M, Morris C. Using EMR-enabled computerized decision support systems to reduce prescribing of potentially inappropriate medications: a narrative review. Ther Adv Drug Saf 2018;9(9):559-573. [doi: 10.1177/2042098618784809] [Medline: 30181862]
12. Anrys P, Boland B, Degryse JM, De Lepeleire J, Petrovic M, Marien S, et al. STOPP/START version 2-development of software applications: Easier said than done? Age Ageing 2016;45(5):589-592. [doi: 10.1093/ageing/afw114] [Medline: 27496930]
13. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, et al. The future landscape of large language models in medicine. Commun Med 2023;3(1):141 [FREE Full text] [doi: 10.1038/s43856-023-00370-1] [Medline: 37816837]
14. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023;9:e45312 [FREE Full text] [doi: 10.2196/45312] [Medline: 36753318]
15. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLOS Digit Health 2023;2(2):e0000198. [doi: 10.1371/journal.pdig.0000198] [Medline: 36812645]
16. Savage T, Nayak A, Gallo R, Rangan E, Chen JH. Diagnostic reasoning prompts reveal the potential for large language model interpretability in medicine. NPJ Digit Med 2024;7(1):20 [FREE Full text] [doi: 10.1038/s41746-024-01010-1] [Medline: 38267608]
17. Goel A, Gueta A, Gilon O, Liu C, Erell S, Nguyen LH, et al. LLMs accelerate annotation for medical information extraction. Proc Mach Learn Res 2023:82-100 [FREE Full text]
18. Agrawal M, Hegselmann S, Lang H, Kim Y, Sontag D. Large language models are few-shot clinical information extractors. In: Goldberg Y, Kozareva Z, Zhang Y, editors. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics; 2022:1998-2002.
19. Kanjee Z, Crowe B, Rodman A. Accuracy of a generative artificial intelligence model in a complex diagnostic challenge. JAMA 2023;330(1):78-80 [FREE Full text] [doi: 10.1001/jama.2023.8288] [Medline: 37318797]
20. SAEM24 Abstracts. Academic emergency medicine. Journal Article 2024;31(S1):8-401. [doi: 10.1111/acem.14906]
21. Safranek CW, Huang T, Wright DS, Wright CX, Socrates V, Sangal RB, et al. Automated HEART score determination via ChatGPT: honing a framework for iterative prompt development. J Am Coll Emerg Physicians Open 2024;5(2):e13133 [FREE Full text] [doi: 10.1002/emp2.13133] [Medline: 38481520]
22. Liu N, Lin K, Hewitt J, Paranjape A, Bevilacqua M, Petroni F, et al. Lost in the middle: how language models use long contexts. Trans Assoc Comput Linguist 2024:157-173. [doi: 10.1162/tacl_a_00638]

23. Shi F, Chen X, Misra K, Scales N, Dohan D, Chi E, et al. Large language models can be easily distracted by irrelevant context. 2023 Presented at: ICML'23: Proceedings of the 40th International Conference on Machine Learning; July 23, 2023; Honolulu, HI.
24. Xiong M, Hu Z, Lu X, Li Y, Fu J, He J, et al. Can LLMs Express Their Uncertainty? An Empirical Evaluation of Confidence Elicitation in LLMs. In: Can LLMs express their uncertainty? An empirical evaluation of confidence elicitation in LLMs. 2024 Presented at: The Twelfth International Conference on Learning Representations; May 7, 2024; Vienna, Austria URL: https://openreview.net/forum?id=gjeQKFxFpZ
25. Leibig C, Allken V, Ayhan MS, Berens P, Wahl S. Leveraging uncertainty information from deep neural networks for disease detection. Sci Rep 2017;7(1):17816 [FREE Full text] [doi: 10.1038/s41598-017-17876-z] [Medline: 29259224]
26. Chua M, Kim D, Choi J, Lee NG, Deshpande V, Schwab J, et al. Tackling prediction uncertainty in machine learning for healthcare. Nat Biomed Eng 2023;7(6):711-718. [doi: 10.1038/s41551-022-00988-x] [Medline: 36581695]
27. Geifman Y, El-Yaniv R. Selective classification for deep neural networks. 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA p. 4885-4894.
28. El-Yaniv R, Wiener Y. On the foundations of noise-free selective classification. J Mach Learn Res 2010;11(53):1605-1641. [doi: 10.5555/1756006.1859904]
29. Liévin V, Hother CE, Motzfeldt AG, Winther O. Can large language models reason about medical questions? Patterns 2024;5(3):100943 [FREE Full text] [doi: 10.1016/j.patter.2024.100943] [Medline: 38487804]
30. Kim S, Lee W. Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? Stat Methods Med Res 2017;26(1):142-154. [doi: 10.1177/0962280214541852] [Medline: 24996898]
31. Seabold S, Perktold J. statsmodels: Econometric and statistical modeling with python. SciPy Proc 2010. [doi: 10.25080/majora-92bf1922-011]
32. Waskom ML. seaborn: Statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]
33. Hunter JD. Matplotlib: A 2D graphics environment. Comput Sci Eng 2007;9(3):90-95. [doi: 10.1109/mcse.2007.55]
34. Gupta S, Basu A, Nievas M, Thomas J, Wolfrath N, Ramamurthi A, et al. PRISM: patient records interpretation for semantic clinical trial matching system using large language models. NPJ Digit Med 2024;7(1):305 [FREE Full text] [doi: 10.1038/s41746-024-01274-7] [Medline: 39468259]
35. Ferber D, Hilgers L, Wiest IC, Leßmann ME, Clusmann J, Neidlinger P, et al. End-to-end clinical trial matching with large language models. ArXiv Preprint posted online on July 18, 2024. [doi: 10.48550/arXiv.2407.13463]
36. Beattie J, Neufeld S, Yang D, Chukwuma C, Gul A, Desai N, et al. Utilizing large language models for enhanced clinical trial matching: a study on automation in patient screening. Cureus 2024;16(5):e60044 [FREE Full text] [doi: 10.7759/cureus.60044] [Medline: 38854210]
37. Nievas M, Basu A, Wang Y, Singh H. Distilling large language models for matching patients to clinical trials. J Am Med Inform Assoc 2024;31(9):1953-1963. [doi: 10.1093/jamia/ocae073] [Medline: 38641416]
38. Omar M, Agbareia R, Glicksberg BS, Nadkarni GN, Klang E. Benchmarking the confidence of large language models in clinical questions. MedRxiv Preprint posted online on September 10, 2024. [doi: 10.1101/2024.08.11.24311810]
39. Savage T, Wang J, Gallo R, Boukil A, Patel V, Safavi-Naini SAA, et al. Large language model uncertainty proxies: discrimination and calibration for medical diagnosis and treatment. J Am Med Inform Assoc 2024:e254. [doi: 10.1101/2024.06.06.24308399]
40. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal A. Human-in-the-loop machine learning: a state of the art. Artif Intell Rev 2023;56(4):3005-3054. [doi: 10.1007/s10462-022-10246-w]
41. Naikar N, Brady A, Moy G, Kwok HW. Designing human-AI systems for complex settings: ideas from distributed, joint, and self-organising perspectives of sociotechnical systems and cognitive work analysis. Ergonomics 2023;66(11):1669-1694 [FREE Full text] [doi: 10.1080/00140139.2023.2281898] [Medline: 38018437]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
ED: emergency department
EHR: electronic health record
GEMS-Rx: Geriatric Emergency Medication Safety Recommendations
IRR: interrater reliability
LLM: large language model
PIM: potentially inappropriate medication
STOPP: Screening Tool of Older People's Prescriptions

Edited by M Mardini; submitted 02.12.24; peer-reviewed by K Lee, I Scott; comments to author 02.01.25; revised version received 16.01.25; accepted 21.02.25; published 11.04.25.

Please cite as:
Socrates V, Wright DS, Huang T, Fereydooni S, Dien C, Chi L, Albano J, Patterson B, Sasidhar Kanaparthy N, Wright CX, Loza A, Chartash D, Iscoe M, Taylor RA
Identifying Deprescribing Opportunities With Large Language Models in Older Adults: Retrospective Cohort Study
JMIR Aging 2025;8:e69504
URL: https://aging.jmir.org/2025/1/e69504
doi:10.2196/69504
PMID:

Exploring Older Adult's Views of the Age-Inclusivity of Physical Activity Websites Using the Think Aloud Method: Qualitative Analysis

Veda Clemson[1], BSc, MSc; Elisabeth Grey[2,3], BSc, MSc, PhD; Julie Barnett[4], BSc, PhD; Ella Burfitt[1], BSc; Fiona Gillison[1], BSc, MSc, PhD

[1]Department for Health, University of Bath, Claverton Down, Bath, United Kingdom

[2]The National Institute for Health and Care Research Applied Research Collaboration West (NIHR ARC West), University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, United Kingdom

[3]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

[4]Department of Psychology, University of Bath, Bath, United Kingdom

Corresponding Author:
Fiona Gillison, BSc, MSc, PhD
Department for Health, University of Bath, Claverton Down, Bath, United Kingdom

Abstract

Background: Older adults are the least active in our society and may face additional barriers to taking part in physical activity compared with those experienced by younger people because of factors such as lower digital literacy and negative stereotypes of aging.

Objective: This study aimed to explore how older adults navigate websites that provide access to physical activity opportunities and facilities and make judgments about their suitability.

Methods: Semistructured interviews were embedded within a think-aloud approach. Participants were shown a series of websites and asked to navigate through the websites as if they were going to take up what was on offer, articulating their thoughts and comments out loud as they progressed. Participants viewed up to 4 websites, rotated from a pool of 8, including leisure centers, exercise products, gyms, or community organizations. Additional questions were asked about perceptions of the inclusivity of the websites at the end of the interview. Digital recordings were made and transcribed verbatim, and analyzed using thematic analysis.

Results: Nineteen participants (6 male and 13 female) aged between 65 and 84 years were recruited from southern England; one-third reported having poor digital ability prior to taking part. Three overarching themes relating to the research question were identified as follows: (1) signals of age-inclusivity, (2) limiting beliefs, and (3) confidence in making judgments. Older adults inferred a lot of information about how welcome they would be in physical activity settings from the images and language used on websites. They showed a preference for imagery that was inclusive of age, body shape, and physical ability, not only for those depicting older adults themselves. Some adults reported firm views about the type and intensity of physical activity that is appropriate for older adults, and many expressed a specific dislike of gyms, based on both the (young) age of most users and perceived emphasis towards aesthetic rather than health-related exercise. While most participants could navigate websites successfully, they preferred to visit venues and speak to staff to gain greater confidence that they would feel welcome and that the activities would be at a suitable level.

Conclusions: Websites providing access to physical activity could be more inclusive of older adults by using more diverse imagery, providing clearer descriptions of the activities on offer, along with details of the level of fitness or ability needed to take part, and providing alternatives to web-based booking. Additional societal-level approaches to reducing age-limiting self-stereotyping may also be useful in expanding the opportunities for older adults to access mainstream provisions for physical activity.

(JMIR Aging 2025;8:e68951) doi:10.2196/68951

KEYWORDS

ageism; inclusivity; digital literacy; older adult; physical activity

Introduction

Older adults are the least active in our society [1], yet they could derive significant benefits from physical activity in terms of preventing and managing chronic disease, promoting strength and function [2], and reducing morbidity [3]. The UK government recommends 150 minutes of moderate-intensity activity per week for older adults, as well as breaking up sedentary periods and including exercise that improves muscle strength, balance, and flexibility [4,5]. Data for 2022 to 2023 from Sport England estimate that around 53% of people older than 65 years achieve this, with 36% achieving less than 30 minutes of activity per week [6]. Activity in later life reduces the risk of disability, the development of noncommunicable diseases, and the decline in functional capacity associated with aging [7]. There may also be benefits to cognitive health [8].

Older adults experience many of the same barriers reported by younger adults to physical activity, including time, cost, and enjoyment, and lack of confidence or fitness [9,10]. Additional barriers that are more commonly seen in older adulthood relate to physical function, such as ill health or disability, safety concerns, and fear of falling [10]. However, there are also indirect barriers that may influence engagement with facilities and activities in older age, including poorer digital literacy, digital exclusion [11], and negative stereotypes around aging [12]. Considering digital literacy, while many older adults do own smartphones or other digital devices, digital acceptability and access decline with age [13]. The physical activity industry mirrors the rest of society in having become more digitized, for example, through moving to web-based bookings, payments, and advertising, and the development of increasingly technical exercise equipment. These changes have an unintended negative impact on access for older adults.

Older adults face several age stereotypes relevant to the physical activity domain, often being portrayed as frail, easily fatigued, and physically inactive [14]. Some of this stereotyping can be observed in how older adults are treated by others, for example, fitness professionals commonly have (and communicate to greater or lesser degrees) low expectations of older adults' capabilities, which in turn discourages older adults from taking part [15]. Providing positive examples of the capabilities of older adults that present a very high level of ability (ie, exceptional fitness) can also be problematic to some older adults who feel they may not be able to match up to expectations, putting them off taking part [16]. Age stereotypes for physical activity are not only held by younger adults, older adults themselves often share these negative perceptions, for example, feeling that their aging body means they are less able to exercise, and that people their age are out of place in physical activity settings, such as gyms and health clubs [17].

The theory of stereotype embodiment sets out one way in which self-stereotyping may develop, proposing that people internalize negative stereotypes throughout their lives, developing negative attitudes toward older adults, which are directed toward themselves as they age [18]. This leads to unconscious effects on behavior, particularly in settings where age stereotypes are primed (eg, depictions of older adults with limited ability), impacting health behaviors [19]. Little research from this perspective has been undertaken specifically in relation to physical activity, and research exploring the impact of age stereotyping on physical activity from other theoretical perspectives has resulted in some conflicting results. For example, in response to age-based stereotype threats (ie, pressure to avoid confirming or exacerbating stereotypes when taking part in an activity; [20]), some studies show poorer performance in physical activities, such as walking speed [21], while others show improved performance, for example in fitness training [22]. Thus, research to provide greater insight into older adults' responses to how they are portrayed during their first approach to physical activity settings is warranted and could help us to understand the variation in effects.

The aim of this study was to explore how websites providing information about physical activity opportunities may influence the barriers or facilitators of physical activity for older adults. Using a "think aloud" approach, our objectives were to (1) explore the social signals older adults identified of inclusivity, fit, or welcome, and (2) capture their views about the digital capability required to find relevant information. Together, this would provide information on how older adults experience web-based gateways to access physical activity opportunities and what information they draw on in assessing whether the activity on offer was suitable for them.

Methods

Ethical Considerations

Ethics approval was granted by the Research Ethics Committee for Health at the University of Bath (EP 20/21 105). All participants provided informed written consent, and any identifying information was anonymized prior to data analysis. Interviews were conducted between January 2022 and August 2023.

Design

This study was situated within a constructivist research paradigm, selecting a "think aloud" approach in order to explore and understand people's subjective experiences and focus on meaning-making in real time [23]. The think aloud approach has shown previous utility in exploring human-computer interactions and decision-making [24] and the usability of websites in health settings [25].

The approach encourages participants to verbalize their thoughts or feelings while navigating web-based resources. In our study, this was embedded within semistructured interviews to ensure interviewees had been asked comparable questions when interpreting the findings, and to allow participants to express their thoughts and experiences at a more general level [26]. The interviews were all conducted in person by slim, White, young women (aged 18 - 23 years); we include this detail as it was commented on by some participants, so it may have influenced some of their responses.

Participants

For this study, we considered adults to be aged 65 years or older, in line with the UK Office for National Statistics [27]. The

inclusion criteria were individuals aged 65 years and older, who can read and speak fluent English, self-assessed as physically able to, and interested in taking part in physical activity. Participants were recruited through advertisements posted in local community centers, on local social media platforms (such as "Nextdoor"), and through the authors' existing networks. We aimed to recruit 20 older adults to provide a range of experiences, focusing on advertising in lower socioeconomic areas. Participants received travel expenses and refreshments but no payment for their time.

Procedure

Volunteers were sent a participant information sheet and consent form and invited to set up a time for their interview; these arrangements were typically made via email, but could be by phone or letter if preferred. Interviews were conducted in public spaces of the participants' choice (to ensure they would feel comfortable) using a laptop provided by the researcher. Participants were asked to nominate a café local to them or could choose to come to the University. At the start of the interview, the researcher talked through the process and obtained written consent. The researcher then presented between 1 and 4 example websites in turn, working through the think aloud process with each until the participant wished to stop or the 1 hour had passed; participants were offered the choice whether to continue at the end of each website viewing. The order in which websites were presented was rotated to ensure that the findings related to a variety of website designs and topics.

Participants were encouraged to provide an honest opinion on the websites, add any thoughts which came to mind [28], and were prompted to reflect on what they were doing as they browsed (eg, "why did you click on that?"; Multimedia Appendix 1 contains the full topic guide). In line with a relaxed think aloud approach, the interviewer only interacted with the participant when asking prompts or in response to a participant's request for help with navigating the websites [28]. At the end of browsing websites, participants were asked to evaluate the content overall. One example of the questions asked for this was "were there any aspects of the websites that made you feel included or excluded?".

The pool of example websites was chosen to reflect different aspects of the physical activity industry following public engagement activity with 7 older adults as part of a related study (ie, what they consider as part of the industry), and discussion with an advisor from a national representative body. Our aims were to include both commercial and noncommercial services, sites that advertise equipment as well as facilities, and ensure we featured a range of activities that are accessible to someone starting out (eg, activities that do not require extensive technical skills, and can be started at low intensity). The research team reviewed the final set to ensure we included variations in website designs (Multimedia Appendix 1 lists the specific websites chosen). Two examples were available from each of the following 5 categories: local authority providers of physical activity services; fitness industry providers (eg, generalist gyms that focused on building strength and endurance); fitness industry product providers (eg, personal fitness equipment and clothing); community organizations that encourage physical activity (eg, walking groups); and UK physical activity campaigns. The order in which the websites were presented was rotated.

After the interview, participants provided information on their demographics, usual physical activity level, and self-rated health.

Data Analysis

Interviews were audio-recorded using a digital voice recorder, transcribed verbatim by VC, and anonymized. Analysis was conducted using a 6-step thematic analysis [29], including familiarization, generating initial codes (including a code book), organizing codes into clusters of shared meaning (themes and subthemes), and refining and naming the themes with the wider research team. Two transcripts were coded by VC and FG to refine the codebook, and the remainder were then coded by VC and the themes were developed iteratively.

Website tracking software was used to generate a recording of participants' progress in order to verify which specific pages participants were looking at as they spoke. This information has been added in square brackets within quotes where appropriate.

Results

Sample

Participants (6 male and 13 female) were aged between 65 and 84 years and based in Southern England. Fourteen (74%) participants had postschool qualifications, and the majority of the participants considered themselves to be healthy (95% scored >7 on a 10-point scale), with moderate fitness (most reported a steady or brisk walking pace) (Table 1).

Around a third of participants considered that they had poor digital ability, but the majority were able to navigate the websites without difficulty, and the challenges experienced most came from using unfamiliar hardware (ie, using the researcher's laptop, rather than a usual PC at home) rather than difficulty using common website functions (scrolling down, search function, clicking through different pages). At the extremes, 2 participants were not confident in navigating websites unaided, so they asked the researcher to do this for them (P11 and P18), while one participant was a website designer prior to retiring (P4).

Three overarching themes relating to the research question were identified as follows: (1) signals of age-inclusivity, reflecting what participants read into the images and language used on websites; (2) limiting beliefs, reflecting how participants related beliefs about their own physical capability, and age stereotypes to website content; and (3) trust, reflecting older adults' reservations about what could be understood from web-based information alone.

Table . Participant characteristics.

	Age group (years)	Gender	Ethnic group	Highest educational attainment	Usual walking pace	Health rating[a]
1	65 - 69	Man	White	Bachelor degree	Steady	8
2	80 - 84	Woman	White	Bachelor degree	Brisk	7
3	75 - 79	Woman	White	Master or postgraduate qualification	Brisk	7
4	65 - 69	Woman	Black or African	GCSE/O level[b]	Slow	3
5	75 - 79	Woman	White	Master or postgraduate qualification	Steady	8
6	70 - 74	Man	White	GCSE/O level	Brisk	8
7	65 - 69	Man	White	Bachelor degree	Steady	7
8	70 - 74	Man	White	Master or postgraduate qualification	Steady	8
9	75 - 79	Man	White	Master or postgraduate qualification	Steady	7
10	65 - 69	Woman	White	Master or postgraduate qualification	Steady	7
11	65 - 69	Woman	White	Bachelor degree	Brisk	8
12	65 - 69	Woman	White	Bachelor degree	Steady	9
13	75 - 79	Woman	White	Master or postgraduate qualification	Slow	10
14	65 - 69	Woman	White	Master or postgraduate qualification	Steady	9
15	65 - 69	Woman	White	No formal qualifications	Slow	8
16	65 - 69	Woman	White	Master or postgraduate qualification	Brisk	8
17	80 - 84	Woman	White	A Level or equivalent	Steady	7
18	75 - 79	Woman	Asian or Asian British	No formal qualifications	Brisk	7
19	65 - 69	Man	Black or African	Master or postgraduate qualification	Slow	8

[a]Self-rated on a scale of 1 - 10 where 1 is very bad health and 10 is excellent health.

[b]GCSE/O level: General Certificate of Secondary Education Ordinary Level.

Theme 1: Signals of Age-Inclusivity

Images

Participants drew on images, language, and the type of physical activity or service offered in judging whether websites were relevant to them. Images were often the first source of information that participants considered ("I noticed almost immediately there was a mix you know, that in photos there were older people" [P9]) and were often taken very literally as an indication of who uses a particular facility. Participants extrapolated from who was depicted as to whether they would be welcome as a participant:

> *This is the gym… oh… group exercise... oh let's see if there's any for old people. [inspects the photograph]. There's a young lady… that's not for me! Absolutely not for me. If they had a photograph of a group with a couple of older people involved in the same group yep, I don't mind being with young people at all, except I wouldn't like to be the only one.* [P2]

> *It's modelled by people, young people without wobbly bits.* [P15]

While seeing older adults depicted was interpreted as a clear indicator that a website was offering something intended for them, most participants used a broader benchmark and judged websites as age-inclusive when they saw images of people of different ages, genders, races, and abilities, as well as people with disabilities. For example, on articulating why certain websites looked age-inclusive, participants stated:

> *[Because] they're all normal people, as in, of an age that I can relate to. So it's not, it's not big punchy*

> *"do this to build your muscle" type thing it's all about… it's all about keeping active and keeping healthy through, through activity, which is good.* [P6]

> *That one is, I think that's very, the seems quite welcoming that, I'd have a look at it, I would if I saw that picture... Because it's a family there, and there's a child there, it is not as though someone's just exercising and looks really fit (laughs). I mean, I'm not saying they don't look fit but I mean it's not a real over-fit.* [P17]

In addition to images, participants responded positively to clear statements that all ages are welcome:

> *Um… they [community rowing club] particularly stress the headline [ages] 13 to 100 so, it definitely was not focused just on the… you know the normally you associate the, I suppose I associate rowing with the boat race, young fit men or women from um from the Universities etc. That's my first reaction so um this seemed quite inclusive in that sense, quite friendly, seems quite informative.* [P9]

Similarly, images of people of all body sizes were considered positive and inclusive, regardless of their age. As such, body size appeared to serve as a proxy for participants believing an activity would be achievable for them:

> *But I mean, but you see that's still good to get people going isn't it because they're all shapes and sizes [in the image] and I'd rather do classes with people like that then just really old people so that's quite nice* [P17]

> *They are… women and they are older women, kind of thing that they are not… slim ones so they really appeal to the likes of me that it can be me, joining them, you see.* [P7]

While wanting to see older adults depicted, participants could be very particular in how old, or how active the older adults were. Some interviewees distanced themselves from websites where older adults were depicted as being too old or inactive:

> *They are very old people and they don't appeal to me at all. I couldn't interact with anybody like that. I mean, okay, they might get satisfaction out of it but nah. It's not me, it's just not me.* [P15]

Very few older adults expressed an interest in exercise sessions that are solely available to older adults, preferring to see indicators of inclusivity to the whole community, with older and larger people depicted as part of this. This may reflect (as suggested by the quote from P15) the degree to which participants have actively embraced an older identity.

Language

In some cases, the language used on websites signaled to participants that these services or facilities were not for them. This was particularly evident in websites relating to gyms and exercise classes, which included new "branded" terms such as classes called "SHRED," "Trax," and "Rush". Participants viewed this terminology as an immediate turn-off, and an indicator that such activities were not for them:

> *It's delving into a sort of more modern language that I don't necessarily understand, but that's probably more my problem than anybody else's!* [P6]

> *Oh… what are glute exercises? Oh… glute, did you know about glute muscles? I didn't.* [P13]

> *Shred (laughing)… see I, well I'm just looking I'm thinking right, so I immediately focused in on classes and I'm not going to go and do a class that says shred because um I just know that that is just, you know, I wouldn't do it.* [P12]

Some participants were explicit in describing the lack of explanation of such classes as discriminatory:

> *I don't think for my particular age group that they are informative enough and I find that they are a bit discriminative and they're… ageist. That's it.* [P15]

Theme 2: Limiting Beliefs

Perceived Physical Capability

Most participants recognized and described a deterioration in their physical ability as they had aged and expressed a desire to avoid injuries by attempting to do too much. While some participants had long-term conditions such as arthritis that limited their activity, for many, this reflected a fear of injury based on the perceived limitations of age alone. This caution influenced their responses to websites based on the type of exercise on offer:

> *With the equipment, I'd end up… tearing something or ending up injuring something of my body. So, I've come to the conclusion that you know, probably it's best not to be going into gyms.* [P4]

This concern was also reflected in participants' desire for more information about the intensity of activities than was provided on most websites. This is often related to expected levels of fitness and capability to start engaging in the exercise on offer, reflecting the fact that although they might have been experienced in some activities (ie, not a beginner), they knew they were much less fit than when they last took part. For example, 2 participants commented on the website of a local walking group that was marketed towards older adults, and depicted older adults in their imagery:

> *It would be helpful if there was a bit more information, to give you… the severity of the walk, is it rough terrain or not. Because obviously as you get older, the flatter the walk the more… probability that you'll actually go onto it…* [P6]

> *Older people do tend to walk a little bit slower so you want to know is this going to take me a couple of hours, or is this going to take all day?* [P6]

> *My concern would be walking out for the first time … I'd be worried about lagging behind and upsetting people who are experienced walkers. If they had a 'break me in walk' and then oh you are broken in a bit now so you can do the next one, and the next one.* [P15]

Similar concerns were offered about fitness classes:

> *It [the website] didn't exclude me, it's just I, my body and everything else that comes with it, makes me feel… a bit worried to try to go and lift weights or do some strenuous thing that I know my body can't take… Do I need to be a certain fitness level before? That's something I need to know.* [P4]

Others expressed concerns about potentially embarrassing themselves in mixed settings or pushing themselves too far. For example, on looking at a strength training class video, one participant commented:

> *If you're not fit I don't think you would ever look [to go] there because you would feel such a fool. You would, you would think I'm a wimp, I couldn't cope with that, so I wouldn't go.* [P17]

Some older adults also expressed concern that younger instructors or trainers would not be able to adapt to what was needed for older adults, and thus that what was demanded of them would be too great. These participants tended to use the age of instructors depicted on websites as a signal of whether a service or facility was appropriate for them:

> *I don't know whether I've just got more respect in general for older people because they've got more experience, maybe that, I just don't quite believe that someone – sorry about this [referring to the young age of the interviewer] – does a three-year course in whatever, fitness, can necessarily have all the information they need to help older people.* [P11]

> *[I'd need] way more support, but who's going to be supporting me, I mean, some squeaky chicky ain't going to come over and say 'Alright, old love let's sort you out!'* [P15]

Preference for older instructors was not a universally held view, and those who had more experience exercising in mixed-age groups provided examples of younger instructors who adapted classes well. They also flagged the need to differentiate people by ability rather than age:

> *My Pilates class is basically for over 50's but they want to stop it too because they say people don't like being defined by their age… so it's like how do you find another way of saying over 55 without patronising somebody?... But at the same time sometimes you don't know what you're looking for. There's some people who are over 65 and they don't need to look at an over 65 sort of over 50s class because they're really fit and they can still do stuff in an ordinary class.* [P12]

Stereotyping

While participants were generally pragmatic and objective in acknowledging age-related changes in physical capability, there was also some evidence of stereotypical thinking. Some of this appeared to relate to self-stereotyping of what types of physical activity older adults should be doing. The example below relates to a comment made while reading through a list of exercise classes at a municipal leisure center.

> *... I'd go for body balance but yoga, no, spin, sounds too fast, those aren't right for old people… If its high intensity, we [older adults] shouldn't be doing things high intensity.* [P2]

However, many of the interview excerpts coded into this theme related to gyms. It was often hard to disentangle participants' pre-existing attitudes towards gyms as a type of business (and their experience of them being predominantly used by young people), from the specific aspects of the website presented ("I'm actually a little bit against gyms anyway, because, it's a modern thing" [P7]). These views typically reflected participants' attitudes towards gyms rather than what was presented on a website, as well as their experience of gyms.

> *I didn't like going in the actual gym bit because it was all people a lot younger than me and they were all trying to outdo each other.* [P15]

> *I see these [gyms] and I only tend to see young people. You know, cool looking sort of young people going into (particular gym), and that doesn't, that doesn't do anything - this website - to disabuse me of that viewpoint. It comes across as a younger, commercial first [place].* [P9]

> *I mean I wouldn't go anywhere near that place [gym]. You know I, not being rude, but I mean cause to me that is just full of… people who are like nearly old enough to be my grandchild so I'm like just wouldn't (laughing), so I wouldn't go in there no.* [P12]

In some cases, the stereotypical thinking reflected assumptions that gyms are a place people go to do image—rather than health—enhancing activity. Often, the negative statements about young people are conflated with assumptions about why people use gyms, given the greater prevalence of young people using them.

> *It's all a bit showy-off-y to me [going to the gym], rather than, I suspect… intended just to get you fit and keep you fit. It's a bit, it's all a bit show-y.* [P7]

> *When I have been into the odd hotel gym there's generally some very, very narcissistic p- you know people there looking at themselves more in the mirror than they are [exercising], you know what I mean, it er, to a younger person you know a, beautiful blonde there or really fit young bloke you know, you know doing whatever I'm thinking, yep, you don't need that.* [P9]

Some participants were less critical of the reasons why people attend gyms, but in believing gyms are for building muscle, and that building muscle is something that only younger people do, they assumed that gyms were not for them. This was reinforced by images of young fit adults lifting weights:

> *When you get over 65, you're more interested probably in er… you're not going to build muscle, you're not going to, you know, attain any major goals. You want to, keep mobile keep flexible, and if there were more things to build upon flexibility, um… that would be more appealing to people of an older age*

group, yes… There doesn't seem to be any of that on this website. [P6]

They [images] all look quite frightening to me, because they all look so much younger and fitter! … I mean look at the, these guys are all pretty stacked, so. I'll never look like that. Not that I ever did! [P6]

Theme 3: Confidence in Making Judgments

Participants preferred websites where information was presented clearly on the first page, where the font was simple, without too many distractions from moving images. The majority considered information about the proximity of facilities ("See, straight away I think what I want to do is find out, where's the nearest one for me?" [P6]), costs, and links to a person they could contact to be of primary importance. They wanted to be able to see this information without needing to click through multiple screens.

I want to know what the membership is and it's giving me "hear from our members." I want to know where it is, how much it is, swimming pay as you go [P15]

So this is, they're not going to give me any information about courses until… until I open the portal and I'm not interested in that. [P5]

Even among those comfortable with using websites and digital tools, many participants expressed a preference to establish some personal contact before committing; they looked for a phone number of someone to talk to, or stated that they would not book via web, but would use the site to gather information and then visit the venue to book. For some, the need to verify information in person reflected a lack of trust in what is presented via web, rather than poor digital literacy.

It's all a marketing ploy, not really giving me the information that I really want. [P15]

I… can't tell, what they would be like from the website… but it's all part of an advertise-advertising their product um so, so yo-you… that's how I would expect it to be I'd expect it to be advertising a product I wouldn't expect to see any truth there [P7]

This hesitancy may also reflect participants' interest in factors that cannot be ascertained from a website, like the atmosphere of the venue or how they are treated by staff.

If I was going to join the gym, I'd go down there and physically look at it. So, um, but only based on having… drawn up a-a-a-a list of 3 say [from their websites]… and that's what I would do and then go and look at them, and see what I thought and judge it on, on um… the people that are um… are um running it so forth, do you get a welcome er a good welcome there, do they seem to know what they're talking about. [P7]

Hesitancy about using websites to make decisions on the appropriateness of physical activity opportunities appeared to differ from general comfort with digital devices or web-based shopping. For example, P7, who would not trust the "truth" of a leisure center website, was comfortable with buying sports equipment via web ("I buy my sports stuff online, or off mates"). Others reported being able to use a device, but not being comfortable with website bookings as a specific function:

Well I'm not very good at websites to be honest, and actually although I'm on it on my iPad with the gym, my son has to book me online all the time. I can't do it, it's sad init? [P18]

Discussion

Principal Findings

This study explored older adults' reactions and views of the age-inclusivity of websites providing access to physical activity opportunities. Together, the themes showed how older adults drew on the imagery and language of websites to inform their decisions, as well as some existing schemas of what types and intensities of activity are appropriate for older adults (ie, implicit ageism). Participants were encouraged by websites depicting people of mixed ages, body shapes, and abilities using their service, and were not looking for services offered purely for older adults. Many older adults judged whether an opportunity was appropriate for them from factors that signaled the intensity and challenge of the physical activity on offer, motivated by a wish to avoid injury or embarrassment, and felt the ability to keep up with others. Information on the level of ability of fitness required to take part was rarely presented directly, which may have caused participants to rely on images and language to make these inferences. Participants showed a greater preference for and trust in websites that provided clear and up-front information on basic considerations such as cost, location, and contact details. Nonetheless, they reported still being reluctant to believe an opportunity was for them without having visited a venue and spoken to staff.

The think aloud approach has been used in previous research to explore how people use and make sense of websites, and what they base their judgments on [28]. For example, users have been shown to use short case examples to understand the relevance of the site by drawing constant comparisons between the examples presented, their own experience [28]. While in this study few websites that we used included case studies (although these were popular where they did), we did observe a similar effect in how participants were extrapolating information from the individuals shown in the images based on their experience and existing attitudes; for example, some participants inferred knowledge about a "typical user's" motivation and fitness level just from a photo. This process was most closely observed in older adults' responses to images of young people, using this information to form judgments that a facility or service was not for them. In part, this reflected participants' explicit narratives of how their capability and motives for physical activity had changed over time, such that they judged what is suitable for young adults would not be suitable for them. While some settings (eg, gyms [30] and surfing [15]) are already considered a young adults' domain by default, our study showed that in other settings too, seeing images predominantly of young adults was used as a heuristic for assessing an activity as irrelevant or inappropriate. This response was diminished when images included mixed age groups, ethnicities, and body shapes alongside young adults.

Past work has described how older adults who take part in physical activity need to navigate the tension between experiencing the benefits of physical activity, noticing their own physical decline, and dealing with social norms about what type of physical activity is acceptable in older age [30]. This tension appeared to be visible in our study, acting as a lens through which participants were evaluating whether the activities that the website was promoting were relevant and welcoming to them. In line with other work [10], we found that older adults' beliefs about their capabilities were a common barrier to physical activity, and evidence of an apparent acceptance of the inevitability of physical decline [17]. Current health messaging, as well as marketing of physical activity facilities and opportunities, may assume that older adults are interested because they believe that there is scope for improvement, so exposing this alternative perspective—that older adults largely do not feel that improvement in strength and fitness is feasible—could be important to inform future approaches.

While participants in this study showed a good ability to navigate websites and good digital literacy, they still exhibited some lack of confidence in applying this. Two participants either asked for help from the Researcher in navigating the websites or reported that they normally ask a younger family member to make web-based bookings or purchases for them. This aligns with previous research demonstrating that even among those with high levels of competency in using websites, older adults may have lower perceptions of their usefulness [31] and acceptability [11]. In this study, this lower perceived usefulness may relate to the fact that information older adults consider crucial was not provided; websites rarely provided explicit statements about welcoming all ages, contact details, clarity about the level of starting fitness required, or information on how demands can be tailored. However, alternatives to web-based booking, whether by phone or in person, will likely remain an important part of providing an age-inclusive offer, to allow older adults to test out the atmosphere of a venue and sample interactions with staff.

Strengths and Limitations

The use of think aloud interviews worked well in this setting by encouraging participants to verbalize their thoughts and experiences while interacting with websites in "real time." It is argued that this often results in more honest accounts of user experience [32]. A limitation of this approach was that participants were navigating websites that they had not chosen for themselves, on a device that was not familiar to them. Participants may have been less distracted by the mechanics of navigating a website if using a familiar device. Our findings may also have been influenced by the specific websites we chose. While we attempted to mitigate this by identifying a range of different websites and rotating these through interviews, providing choice to participants (eg, allowing them to nominate the types of physical activity there were interested in) may have helped to provide a greater test of participants' views on the inclusivity of the website, rather than the type of activity it offered. In particular, it was hard to disentangle participants' comments on the desirability of gyms and the imagery that accompanied them, reported in theme 2, from their preexisting attitudes.

We attempted to recruit a diverse sample by advertising in areas populated by people of lower income and with greater ethnic diversity, and by seeking support from gatekeepers (eg, local Black and minority ethnic support groups for older adults). However, the majority of those willing to take part were White, highly educated, and living in urban areas, which is a limitation of this study. Other studies conducted with older adults report greater challenges with website navigation [25], which we may have found if we had recruited a sample with a more diverse socioeconomic background. Our participants were also relatively healthy; this is less a limitation and more a reflection of our design, as in order to address our research question, we recruited individuals for whom the websites were relevant, that is, those who are capable of engaging in physical activity. However, we may have discovered different barriers had we included more people with limited physical ability. Related to this, the age of our participants spanned around 2 decades from 65 to 84 years, from those still working to those who have been retired for 20 years. People at either end of this age group, as well as those experiencing greater or lesser physical decline or limitations, may have had different perspectives on who the websites depicted and what was offered. Within this small qualitative study, we were not able to tease out these differences.

Implications and Future Directions

The findings of this study suggest that older adults are more likely to consider that services and facilities are relevant and attractive to them when websites depict a diverse set of users, including an explicit statement indicating people of all ages are welcome, and present clear information to users to judge the level of fitness and ability likely to be required to take part. Remaining physically active for longer is key to reducing age-related disability and disease [7] and promoting good cognitive health [8]. In addition, including older adults in intergenerational settings can foster the mental and physical health benefits associated with social inclusion and connection [33,34]. If increased physical activity participation in the growing population of older adults is to be achieved, their needs need to be accommodated affordably within mainstream facilities and services.

A common reason why older adults wanted more information on the likely demands of the physical activity on offer was to reassure them that they would not be pushed too far, feel embarrassed, and that the activity could be adapted safely for someone of their age and level of ability. Classifications referring to the level of experience (beginner, intermediate, etc) may not be meaningful when someone has historical experience but has not taken part in many years, and a lack of explanation alongside activities that have nontraditional names can lead older adults to infer that this is not intended for them. Evidence of venues using older adults as instructors and coaches, and information to confirm that activities can be adapted to be appropriate for all ability levels and ages (where this is the case) would also increase older adults' confidence in trying out physical activity opportunities. While many older adults can navigate websites competently, they expressed a preference for simple information and were put off by too many moving images and the need to click through multiple links to find information.

Conclusions

This study has provided insight into what older adults notice and look for in websites when exploring opportunities to increase their physical activity. It flagged how their pre-existing beliefs and implicit assumptions may influence whether they consider new options for activity, taking into account perceptions of their own capability, social norms, and some stereotypes around age and motivation for physical activity. These existing beliefs may interact with web content to exacerbate or potentially challenge stereotypes, to encourage older adults to feel more welcome to use mainstream facilities and services that support physical activity.

Acknowledgments

We would like to thank Ms Natanya Cassidy and Ms Abigail Russell for their contribution to the interviews, as well as all the participants who gave up their time to take part. EG's time is supported by the National Institute for Health and Care Research Applied Research Collaboration West (NIHR ARC West)

Authors' Contributions

FG, JB, EG, and EB contributed to the conceptualization of the study. VC, EB, and AR were responsible for data collection. VC and FG conducted the analysis. FG and VC drafted the initial version of the manuscript. All authors contributed to further writing, editing, and refinement of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials containing the interview topic guide and a list of websites used in the study.
[DOCX File, 19 KB - aging_v8i1e68951_app1.docx]

References

1. Lindsay T, Westgate K, Wijndaele K, et al. Descriptive epidemiology of physical activity energy expenditure in UK adults (The Fenland study). Int J Behav Nutr Phys Act 2019 Dec 9;16(1):126. [doi: 10.1186/s12966-019-0882-6] [Medline: 31818302]
2. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]
3. Chudasama YV, Khunti KK, Zaccardi F, et al. Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study. BMC Med 2019 Jun 12;17(1):108. [doi: 10.1186/s12916-019-1339-0] [Medline: 31186007]
4. Sparling PB, Howard BJ, Dunstan DW, Owen N. Recommendations for physical activity in older adults. BMJ 2015 Jan 21;350:h100. [doi: 10.1136/bmj.h100] [Medline: 25608694]
5. Physical activity guidelines: UK chief medical officers' report. : Department of Health and Social Care; 2019.
6. Active lives survey. : Sport England; 2024.
7. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta‐analyses. Scandinavian Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616]
8. Bherer L, Erickson KI, Liu-Ambrose T. A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. J Aging Res 2013;2013(1):657508. [doi: 10.1155/2013/657508] [Medline: 24102028]
9. Bethancourt HJ, Rosenberg DE, Beatty T, Arterburn DE. Barriers to and facilitators of physical activity program use among older adults. Clin Med Res 2014 Sep;12(1-2):10-20. [doi: 10.3121/cmr.2013.1171] [Medline: 24415748]
10. Spiteri K, Broom D, Bekhet AH, de Caro JX, Laventure B, Grafton K. Barriers and motivators of physical activity participation in middle-aged and older-adults - a systematic review. J Aging Phys Act 2019 Sep 1;27(4):929-944. [doi: 10.1123/japa.2018-0343] [Medline: 31141447]
11. Gallistl V, Rohner R, Hengl L, Kolland F. Doing digital exclusion - technology practices of older internet non-users. J Aging Stud 2021 Dec;59:100973. [doi: 10.1016/j.jaging.2021.100973] [Medline: 34794717]
12. Marquet M, Chasteen AL, Plaks JE, Balasubramaniam L. Understanding the mechanisms underlying the effects of negative age stereotypes and perceived age discrimination on older adults' well-being. Aging Ment Health 2019 Dec 2;23(12):1666-1673. [doi: 10.1080/13607863.2018.1514487]
13. Choudrie J, Pheeraphuttranghkoon S, Davari S. The digital divide and older adult population adoption, use and diffusion of mobile phones: a quantitative study. Inf Syst Front 2020 Jun;22(3):673-695. [doi: 10.1007/s10796-018-9875-2]
14. Ng R. Societal age stereotypes in the U.S. and U.K. from a media database of 1.1 billion words. Int J Environ Res Public Health 2021 Aug 21;18(16):8822. [doi: 10.3390/ijerph18168822] [Medline: 34444578]
15. Jin B, Harvey IS. Ageism in the fitness and health industry: a review of the literature. J Aging Phys Act 2021 Feb 1;29(1):99-115. [doi: 10.1123/japa.2019-0230] [Medline: 32759410]

16. Wangler J, Jansky M. Media portrayal of old age and its effects on attitudes in older people: findings from a series of studies. Humanit Soc Sci Commun 2023;10(1):1-9. [doi: 10.1057/s41599-023-01671-5]
17. Evans AB, Sleap M. "You feel like people are looking at you and laughing": older adults' perceptions of aquatic physical activity. J Aging Stud 2012 Dec;26(4):515-526. [doi: 10.1016/j.jaging.2012.07.004] [Medline: 22939548]
18. Levy B. Stereotype embodiment: a psychosocial approach to aging. Curr Dir Psychol Sci 2009 Dec 1;18(6):332-336. [doi: 10.1111/j.1467-8721.2009.01662.x] [Medline: 20802838]
19. Levy BR, Myers LM. Preventive health behaviors influenced by self-perceptions of aging. Prev Med 2004 Sep;39(3):625-629. [doi: 10.1016/j.ypmed.2004.02.029] [Medline: 15313104]
20. Lamont RA, Swift HJ, Abrams D. A review and meta-analysis of age-based stereotype threat: negative stereotypes, not facts, do the damage. Psychol Aging 2015 Mar;30(1):180-193. [doi: 10.1037/a0038586] [Medline: 25621742]
21. Barber SJ, Hamel K, Ketcham C, Lui K, Taylor-Ketcham N. The effects of stereotype threat on older adults' walking performance as a function of task difficulty and resource evaluations. Psychol Aging 2020 Mar;35(2):250-266. [doi: 10.1037/pag0000440] [Medline: 31971413]
22. Deshayes M, Palermo A, Korchi K, Philippe AG. Influence of negative stereotype on physical activity level among older adults during a training session. Front Sports Act Living 2022;4:998724. [doi: 10.3389/fspor.2022.998724] [Medline: 36506719]
23. Koro-Ljungberg M, Douglas EP, Therriault D, Malcolm Z, McNeill N. Reconceptualizing and decentering think-aloud methodology in qualitative research. Qual Res 2013 Dec;13(6):735-753. [doi: 10.1177/1468794112455040]
24. Cotton DRE, Gresty KA. The rhetoric and reality of e‐learning: using the think‐aloud method to evaluate an online resource. Assess Eval High Educ 2007 Oct;32(5):583-600. [doi: 10.1080/02602930601116920]
25. Bolle S, Romijn G, Smets EMA, Loos EF, Kunneman M, van Weert JCM. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Jul 25;18(7):e208. [doi: 10.2196/jmir.5618] [Medline: 27457709]
26. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
27. Profile of the older population living in england and wales in 2021 and changes since 2011. Office for National Statistics. 2025. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/profileoftheolderpopulationlivinginenglandandwalesin2021andchangessince2011/2023-04-03 [accessed 2025-03-17]
28. Breuning M, Schäfer-Fauth L, Lucius-Hoene G, Holmberg C. Connecting one's own illness story to the illness experiences of others on a website-an evaluation study using the think aloud method. Patient Educ Couns 2020 Jan;103(1):199-207. [doi: 10.1016/j.pec.2019.08.014] [Medline: 31466883]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Tulle E, Dorrer N. Back from the brink: ageing, exercise and health in a small gym. Ageing Soc 2012 Oct;32(7):1106-1127. [doi: 10.1017/S0144686X11000742]
31. Campelo AM, Katz L. Older adults' perceptions of the usefulness of technologies for engaging in physical activity: using focus groups to explore physical literacy. Int J Environ Res Public Health 2020 Feb 11;17(4):1144. [doi: 10.3390/ijerph17041144] [Medline: 32053937]
32. Alhadreti O, Mayhew P. Rethinking thinking aloud: a comparison of three think-aloud protocols. Presented at: CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; Apr 21-26, 2018; Montreal, Canada. [doi: 10.1145/3173574.3173618]
33. Krzeczkowska A, Spalding DM, McGeown WJ, Gow AJ, Carlson MC, Nicholls LAB. A systematic review of the impacts of intergenerational engagement on older adults' cognitive, social, and health outcomes. Ageing Res Rev 2021 Nov;71:101400. [doi: 10.1016/j.arr.2021.101400] [Medline: 34237435]
34. Zhou F, Zhang H, Wang HY, Liu LF, Zhang XG. Barriers and facilitators to older adult participation in intergenerational physical activity program: a systematic review. Aging Clin Exp Res 2024 Feb 12;36(1):39. [doi: 10.1007/s40520-023-02652-z] [Medline: 38345651]

Edited by Y Jiang; submitted 27.11.24; peer-reviewed by A Chasteen, D Sullivan, M Griffin; revised version received 17.03.25; accepted 01.05.25; published 19.06.25.

Please cite as:
Clemson V, Grey E, Barnett J, Burfitt E, Gillison F
Exploring Older Adult's Views of the Age-Inclusivity of Physical Activity Websites Using the Think Aloud Method: Qualitative Analysis
JMIR Aging 2025;8:e68951
URL: https://aging.jmir.org/2025/1/e68951
doi:10.2196/68951

Original Paper

Factors Influencing Older Adults' Perception of the Age-Friendliness of Their Environment and the Impact of Loneliness, Technology Use, and Mobility: Quantitative Analysis

Eric Balki[1], BSc, MA, MSc, PhD; Niall Hayes[2], BSc, MSc, PhD; Carol Holland[1], BSc, PhD

[1]Department of Health Research, Faculty of Health and Medicine, Lancaster University, Lancaster, United Kingdom

[2]Keyworth Building, Leeds University, Leeds, United Kingdom

Corresponding Author:
Eric Balki, BSc, MA, MSc, PhD
Department of Health Research
Faculty of Health and Medicine
Lancaster University
Health Innovation One, Sir John Fisher Dr
Bailrigg
Lancaster, LA1 4AT
United Kingdom
Phone: 44 1524 65201
Email: e.balki@lancaster.ac.uk

Abstract

Background: The World Health Organization's (WHO) publication on age-friendly environments (AFEs) imagines future cities to become more age-friendly to harness the latent potential of older adults, especially those who have restricted mobility. AFE has important implications for older adults in maintaining social connections, independence, and successful aging-in-place. However, technology is notably absent in the 8 intersecting domains of AFEs that the WHO imagines improve older adult well-being, and we investigated whether technology should form a ninth domain. While mobility was severely restricted, the COVID-19 pandemic provided an opportunity to test how older adults' perceptions of their AFE changed and what role technology was playing.

Objective: This study examined how life-space mobility (LSM), a concept for assessing patterns of functional mobility over time, and loneliness impacted perceived AFEs and the moderating effect of technology. It also explores whether technology should play a greater role as the ninth domain of the WHO's imagination of the AFE of the future.

Methods: In this cross-sectional quantitative observation study, data from 92 older adults aged 65-89 years were collected in England from March 2020 to June 2021 during the COVID-19 pandemic. The Life-space Questionnaire, Technology Experience Questionnaire, UCLA (University of California, Los Angeles) Loneliness Scale, and age-friendly environment assessment tool were used. Correlation and moderation analyses were used to investigate relationships between variables.

Results: Most participants (86/92, 93%) had not left their immediate town in the previous 4 weeks before the interview. Restricted LSM was positively correlated to the age-friendly environment assessment tool, that is, rising physical isolation was linked to a better perception of AFEs; however, we discovered this result was due to the moderating impact of increased use of technology, and that restricted LSM actually had a negative effect on AFEs. Loneliness was correlated negatively with the perception of AFEs, but technology use was found to moderate the impact of loneliness.

Conclusions: Pandemic-related LSM restrictions impacted perceived AFEs and loneliness negatively, but technology played a moderating role. The findings demonstrate that technology could be considered as a ninth domain in the WHO's assessment of AFEs for older adults and that there is a need for its explicit acknowledgment.

(JMIR Aging 2025;8:e67242) doi:10.2196/67242

KEYWORDS

COVID-19; age-friendliness of environments; physical isolation; digital communication technologies; loneliness; cross-sectional; WHO; World Health Organization; older adults; reduced mobility; age friendliness of environments; adult well-being; social

connections; aging in place; life-space mobility; LSE; functional mobility; UCLA loneliness scale; age-friendly environment assessment tool; AFEAT

Introduction

Background

Enabling the development of age-friendly environments (AFEs), defined as a physical and social setting that helps people age well and participate in their communities to promote "aging-in-place" for the mental and physical well-being of older adults, has become an increasingly important policy issue. This is in part a response to an aging population dynamic, urbanization, intensification of housing concerns, and community preferences, causing an increase in the deterioration of mental wellness, anxiety, stress, and depression among other disorders [1]. Older adults' perception of AFEs directly affects their quality of life (QOL) and known predictors of depression, including loneliness [2,3]. In urban environments, older adults tend to spend much of their time in their local neighborhoods and are therefore sensitive to change [4,5]. Environmental degradation, such as lack of resources, restricted access to health care, or crime, brings additional challenges.

Previous studies have demonstrated that the perception of AFEs affects the action space of older people, affecting their social participation [6]. It is known to positively moderate the relationship between frailty, mental well-being, and depression, particularly in older adults with poor health, limited mobility, and cognitive decline [1,7,8]. Moreover, a lack of social opportunities can result in loneliness, social isolation, and worsened depression, all of which exacerbate cognitive decline, QOL, and frailty [8-14]. Aboderin et al [15] and Ng et al [16] also noted a reduction in older adults' psychological resilience due to sustained stress and anxiety resulting from crowding, lack of space, and other issues that revolve around a poor age-centered environment. Other research has demonstrated a correlation between loneliness and AFEs [17-21], and identified factors linking the psychological health of economically disadvantaged older adults with their perception of AFE. Rantakokko et al [22] described the impact of person-environment interaction on mental well-being, demonstrating a mechanism by which older adults who experienced loneliness perceived obstacles with their environment. Stafford [23] demonstrated that older age insularity and the accompanying withdrawal from social interaction can result in deteriorating relationships, loneliness, poorer mental health, and a diminished perception of AFEs. If the environment is not conducive to aging-in-place, it causes difficulties in accessing services [24] and an increased risk of physical and mental health care needs [25]. As such, the perception of AFEs must be considered carefully to improve understanding of older adults aging-in-place and mental well-being.

Life-space mobility (LSM) describes the physical environment a person inhabits daily, structured into various zones (called life-space zones), centered on an anchor (eg, bedroom), and expanding outward into the rest of the house, house perimeter, local community, neighborhood, or town [26,27]. This concept corresponds to individuals' functional mobility enabling meaningful participation in community activities. It reflects how older adults move across life-space zones over a given period while incorporating frequency and independence.

According to the United Nations Development Program report "The Sustainable Development Goals and COVID-19" [28], there are unprecedented ongoing burdens on mental health for older adults. Social distancing orders during the COVID-19 pandemic were conceived as advantageous to protect potentially vulnerable populations, such as older adults [29], from disease transmission; however, it has led to long-lasting effects on mental well-being.

Restricted LSM impedes older adults' access to their [30-32] choice of environments and is associated with potential adverse mental health outcomes [27]. In addition, previous research has demonstrated that LSM restrictions reduce participation in out-of-home [33] activities and negatively affect QOL [9,34], which is in turn associated with loneliness [35].

Although pandemics create restricted mobility [36], older adults with knowledge of technology can capitalize on a digitally enabled AFE to avoid loneliness and LSM restrictions that create a strain on mental well-being. It enables older adults to engage in internet-based activities they may have enjoyed in person and gain access to health care, civic services, therapy, counseling [37], and resources that help with mental well-being [38,39]. We therefore expected both variables (LSM and loneliness) to impact AFEs negatively and for technology to moderate their impact.

Feldman and Oberlink [40] pioneered the concept of AFEs, identifying important components such as social engagement, enhancing independence, and optimizing mental functioning and well-being. Age-friendliness is derived from an ecological model of aging, in which a person's mental well-being results from the interaction between their functional and cognitive capacity (competencies) and the environmental characteristics that exert pressure on these competencies (environmental stress) [41].

The age-friendly environment assessment tool (AFEAT) was designed to assess older adults' perception of their environment and focuses on individual-oriented age-friendliness and individual-environment interaction, providing a more holistic picture [11,42,43]. Self-perceived AFEs are associated with improved QOL and mental health regardless of older adult frailty or abilities [8].

Although there is currently no universally accepted definition of an AFE, the World Health Organization (WHO) tentatively defines such communities as those where "policies, services, settings and structures support and enable people to age actively." The WHO [1] publication imagines future cities to become more age-friendly to harness the latent potential of older adults through 8 intersecting domains addressing obstacles to older adults' mental well-being: respect or social inclusion, outdoor spaces and buildings, housing, social participation, transportation, communication and information, civic

participation and employment, and community support and health services. Explicit mention of technology is notably absent.

The WHO framework has been criticized for the exiguous technology element in the domains, prompting a reexamination [44]. Incorporating technology into AFEs in its broadest sense has become an increasingly important area for aging independence and mental well-being [45] in recent years.

Marston and van Hoof [44] discussed the incorporation and use of technology within the AFE assessment agenda, and a recent systematic umbrella review elaborated on the advantages of technology interventions for social connectedness [46]. Their call for the inclusion of digital technology as a domain in AFE evaluations is beginning to resonate.

For example, Pedell et al [47,48] advocated for digital elements to encompass all aspects of environmental age-friendliness, in addition to mental and physical aspects, to realize benefits for older adults. Moreover, Reuter et al [49] acknowledged the WHO's age-friendly city initiative, considering an aging population amid increasing urbanization. However, they determined that such initiatives overlooked technology as a critical component of global digitalization.

Research Objectives

Our study aimed to examine the potentially complex relationship between internal (loneliness) and external (LSM) factors that influence older people's perception of AFEs and to determine whether technology moderates their impact. We aimed to answer a question with an important implication: should technology be included as a ninth domain in the WHO's global age-friendly cities guide in the assessment of AFEs in cities and communities for older adult mental well-being?

We set out to test the following hypotheses:

- Hypothesis (H1): LSM restriction is negatively correlated with perception of AFEs.
- H2: Increased feelings of loneliness are correlated with poor perception of AFEs.
- H3: Technology use moderates the impact of LSM restrictions on the perception of AFEs.
- H4: Technology use moderates the impact of loneliness on the perception of AFEs after considering the LSM restriction effect.

Methods

Study Design

This was a cross-sectional quantitative observational design. This report follows the STROBE (Strengthening Reporting of Observational Studies in Epidemiology) guidelines [50] (STROBE checklist mentioned in Multimedia Appendix 1).

Setting

Participants were recruited from the United Kingdom. Data were collected from January 16, 2020, to June 21, 2021, a period when social distancing mandates were enforced and social engagement outside the home was restricted.

Participants

Eligible participants were required to be living in their own home, be proficient in English, and be 65 years or older. Older people who lived in nursing or care homes or those with cognitive decline or mental health issues were excluded. Volunteers were recruited via advertisements posted in resource centers for older adults, housing associations, third-sector organizations, social activity clubs, local senior groups, direct human interaction, and word-of-mouth recommendations. Volunteers were instructed to either call and leave a voicemail or send an email indicating their willingness to participate; a return call confirmed their eligibility.

G*Power software (Version 3.1; Erdfelder, Faul, and Buchner) was used to calculate the minimum sample size required for the empirical validation of the tested moderation model. Multiple regression was used, with effect size f^2 of 0.15, power of 0.80, and 3 predictors. The recommended sample size was determined to be 87. A total of 110 participants enrolled; however, 18 did not complete all questionnaires and were excluded. The sample achieved included 92 people between the ages of 65 and 92 years (mean age 74.6, SD 7.23 years). All participants identified as either male or female, with more females (55/92, 60%) than males. More than 89% of the participants were White, with less than 11% of minority ethnicities (7 British Asian, 3 British Black). Having collated various demographic information such as age, gender, ethnicity, and education level, we were able to ascertain that participants emanated from diverse sociodemographic backgrounds.

Variables and Measures

All participants filled out a health history questionnaire based on the Scientific Advisory Group for Emergencies (SAGE) Encyclopedia of Communication Research Methods [51].

Loneliness was measured using the 20-item UCLA (University of California, Los Angeles) Loneliness Scale [52], with scores ranging from 20 to 80. Higher scores reflected higher loneliness (Cronbach α=0.88).

Utilization of technology was evaluated using the Technology Experience Questionnaire [53]. The participants were given a list of technologies (communication, computer, daily, health, recreational, and transportation technology) and asked to rate their familiarity with and use of each on a 5-point scale. Scores ranged from 0 to 180, with higher scores indicating greater use and familiarity with technology (Cronbach α=0.84).

The perception of the AFE was assessed by the AFEAT. This is a 10-item measure that uses a 5-point Likert scale, scoring items from (1) strongly disagree to (5) strongly agree and gauging participants' perceptions of their home, their local communities, the resources within the environment, and their appropriateness for meeting their daily needs. The scores ranged from 0 to 50, with higher scores representing a more positive perception of the age-friendliness of the environment (Cronbach α=0.75).

LSM was measured using the Life-Space Questionnaire [26]. Participants were asked yes or no questions about specific places they visited in the last 4 weeks, starting with another room in

their current residence and increasing the distance to a location outside England. The scores ranged from 9 to 18, with higher scores demonstrating greater restriction of LSM (Cronbach α=0.90).

Procedure

Telephone surveys collected information on loneliness, technology use, LSM, and perceptions of AFEs, in addition to basic demographic information (eg, age, education, gender, and ethnicity). Google Analytics was used to collect and tabulate the data. Participants completed the assessments across 14 months.

Statistical Analysis

Analyses were performed using IBM SPSS version 28, using a minimum significance level of 95% probability. There were no missing data, and participants completed all questions. The variables of AFEAT, loneliness, technology use, and LSM were inspected for kurtosis and skewness to assess their distribution deviations from normality via a histogram with simulated overlapping normal curves. Moreover, the homoscedasticity of the residuals was checked using a standardized residual versus a standardized predicted plot. Using the Mahalanobis (P<.001) and Cook distances, we determined whether high leverage points, significant outliers, or highly influential points exist by examining a scatterplot matrix of the dependent and continuous independent variables. A linear regression was performed to check the included variance caused by the data point and if it needed to be removed from the dataset. The criterion for discarding observations was the inability to meet 2 of the distance measures' 3 gauges. However, no outliers were found that would significantly impact the findings, and thus, none were removed. The confirmation of the independence of the observations and the assumption of no autocorrelation in the residuals was checked using the Durbin-Watson d-statistic.

The initial descriptive analyses contained means, frequencies, and SD. Pearson product-moment correlation coefficients (r) were calculated to determine whether associations exist between the variables. The same correlational analysis was used to determine whether the perception of AFE is correlated with LSM during the pandemic (H1) and whether loneliness is correlated with the perception of AFEs (H2).

Hayes' [54] PROCESS macro for SPSS with model 1 was applied to investigate the moderating effects of technology use on the relationship between LSM and AFE (H3) and technology use on the relationship between loneliness and AFEs. If the standardized coefficients of the interaction terms were significant (P<.05) or marginally significant (P<.09), we conducted a simple slope test to examine the interaction effect at different levels to reveal the nature of significant interactions to further explain the moderating effect.

Ethical Considerations

Participants accessed an information sheet either via email or read on the phone and were allowed to ask questions before giving their consent. Before completing the questionnaires, each participant gave informed consent to volunteer without compensation and participate in the study. All participants were fully anonymized.

All participants were informed of their rights to withdraw at any point in the research and informed about anonymity. The ethical procedures were aligned with the guidelines of the British Psychological Society, and the study received ethical approval from the Lancaster University Faculty of Health and Research Ethics Committee (reference number FHMREC19121). Data captured via telephone first confirmed the participant's identity and was recorded in spreadsheets and anonymized thereafter.

Results

Overview

Table 1 shows mean, SD, kurtosis, and skewness values and Shapiro Wilk test results. Kurtosis and skewness values had a relatively small range of ±1; we determined that the normal distribution deviation was insignificant. The distributions of the variables of loneliness, AFE, technology experience, and LSM were close to normal.

Table 1. Descriptive statistics for loneliness, life-space, age-friendly environment assessment tool, and technology experience (N=92).

	Scores					
	Minimum	Maximum	Mean (SD)	Skewness	Kurtosis	Shapiro Wilk test
UCLA[a] loneliness score	21	80	47.49 (17.814)	0.204	−1.630	0.192[b]
Life-space mobility	10	18	13.70 (1.595)	0.515	0.035	0.201[b]
Age-friendliness of environment	0	35	19.51 (9.687)	0.019	−1.399	0.168[b]
Technology experience	48	175	116.87 (40.951)	−0.260	−1.624	0.204[b]

[a]UCLA: University of California, Los Angeles.

[b]P<.001 under moderate.

Participants demonstrated high levels of loneliness, with 44% of older adults demonstrating loneliness scores greater than 50, with scores above 40-50 considered moderate loneliness and scores greater than 50 considered high [55]. LSM scores were high, with more than 93% of the participants scoring >11, showing that they had not been outside their immediate town. Previous prepandemic studies with similar sample sizes and methodology reported almost half that score [26]. The

perceptions of AFEs were mixed, with a mean score of 19.51 (SD 9.69), demonstrating both positive and negative perceptions.

Prepandemic data from Garner and Holland's [8,56] studies provided a mean of 42.2, taken from 132 participants based in England, indicating a more positive perception of AFEs before the pandemic. Most participants scored above 125 (56%) for technology, demonstrating frequent use and familiarity with technology in general [53].

Next, we calculated the Pearson correlation coefficients to establish the relationships between loneliness, technology, LSM, and AFE perception to test H1 and H2. A correlation matrix of the variables was examined and is presented in Table 2.

Table 2. Correlational analysis between variables (N=92).

Variables	UCLA[a] Loneliness Scale	Life-space mobility	Age-friendliness of environment	Technology experience
UCLA Loneliness Scale				
r	1	−0.483[b]	−0.698[b]	−0.631[b]
P value	—	.002	.006	.003
Life-space mobility				
r	−0.483[b]	1	0.461[b]	0.430[b]
P value	.003	—	.003	.004
Age-friendliness of environment				
r	−0.698[b]	0.461[b]	1	0.667[b]
P value	.003	.004	—	.003
Technology experience				
r	−0.631[b]	0.430[b]	0.667[b]	1
P value	.003	.004	.002	—

[a]UCLA: University of California Los Angeles.

[b]*P*<.01.

[c]Not applicable.

LSM is Negatively Correlated With the Perception of AFE During the Pandemic (H1)

The correlation between LSM and AFE perception was statistically significant (r=0.461, P<.001) but positively correlated, which was contrary to the hypothesis (Table 2). This meant that higher LSM scores associated with restricted mobility were correlated with a greater positive perception of AFEs. Although this rejected H1, it was a notable result.

Loneliness Is Negatively Correlated to a Perception of AFEs (H2)

The correlation between loneliness and AFE perception was statistically significant (r=−0.698, P<.001) and negatively correlated. This meant that greater loneliness was correlated with more negative perceptions of AFEs, thus confirming H2.

Technology Use Moderates the Impact of LSM Restriction on the Perception of AFEs (H3)

Model 1 was used in the PROCESS 4.0 macro for SPSS to examine the moderation effect proposed in H3 [54], as shown in Figure 1, which shows the moderation role played by technology use in the relationship between LSM and AFE perception.

Here, all continuous variables were converted to z scores for use in the model as suggested by Frazier et al [57] and Hayes [54] (ie, via z scoring, expressed as the deviation from their sample means in SD units). The unconditional interaction of LSM and technology use was insignificant (β=0.1921, t_1=1.963; P=.06; Table 3).

All variables in the model are standardized and brought into the regression equation.

Using the Aiken and West [58] method, a simple slope test was used to analyze the conditional effect of technology use between LSM and AFEs (ie, whether technology use moderates the relationship between LSM and AFEs). As illustrated in Figure 2, when technology experience was high, LSM and AFEs were significantly positively correlated (β simple: mean 1, SD 2.2037, t=3.2216; P=.003), indicating that older adults' perception of their environment was more positive when they used technology more.

In contrast, the relationship between LSM and AFEs was not obvious when the technology experience was low (β simple: mean −1, SD −0.3429, t=−0.3510; P=.73). Thus, there appears to be a positive relationship between LSM restriction and perception of AFEs when technology use was high but not when it was low (explaining the unexpected direction of H1). Furthermore, note the slight downward slope indicating the negative impact of LSM restrictions on the perception of AFEs when technology experience was low. This hints that in the absence of technology use, LSM restrictions had a detrimental impact on the perception of AFEs. This confirms H3 and

explains the initial rejection of the unconditional (overall) moderation impact of technology experience on LSM by a small margin (P=.06).

Figure 1. Moderating role of technology experience on the relationship between life-space mobility and age-friendliness of environments.

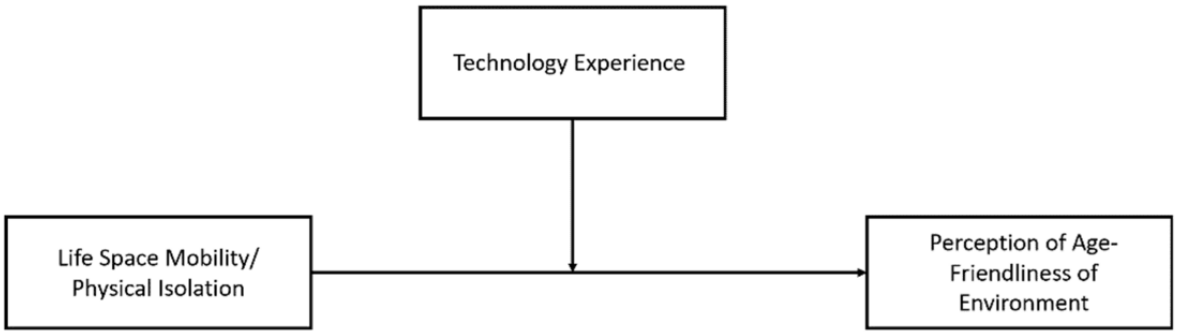
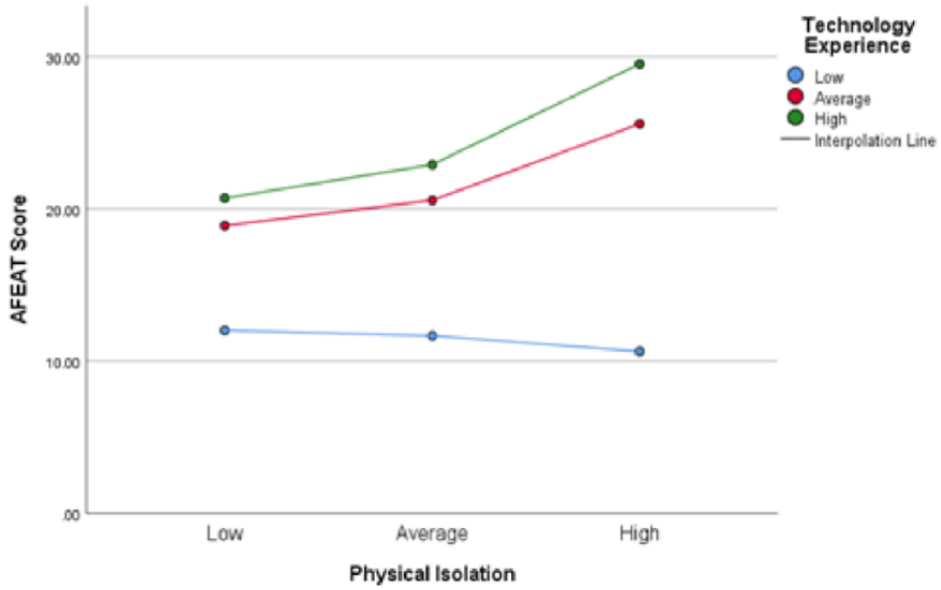


Table 3. Interaction between life-space mobility and technology experience.

	Overall fit indicators			Significance of standardized coefficient	
	R	R^2	*F* test (*df*)	β	*t* test (*df*)
Variables	0.71	0.50	29.7962 (1,90)		
Life-space mobility (ZLS1)				0.1787	2.1002[a] (90)
Technology experience (ZTE1)				0.6155	7.1839[a] (90)
ZLS1*ZTE1				0.1921	1.9628 (90)

[a]P<.001.

Figure 2. Moderating role of technology experience on the relationship between life-space mobility and age-friendly environment (simple slope test). AFEAT: age-friendly environment assessment tool.



Technology Use Moderates the Impact of Loneliness on the Perception of AFEs After Considering the Impact of LSM (H4)

Model 1 was used in the PROCESS 4.0 macro for SPSS to examine the moderation effect of technology use on the relationship between loneliness and perception of AFEs, as proposed in H4 [54] and shown in Figure 3, which shows the moderating effect of technology experience on the relationship between loneliness and the perception of AFEs while controlling for LSM.

All continuous variables were converted to *z* scores. As shown in Table 4, the relationship between loneliness and technology experience was significant (β=−0.3829, t_{90}=−5.1518; P<.001), but the impact of LSM when added to the regression model was not significant (β=0.7151, t_{90}=1.665; P=.07), showing that

although technology use had a moderating impact on the relationship between loneliness and AFE perception, the contribution of LSM was not significant on the model once these other variables had been taken into account.

All variables in the model are standardized and brought into the regression equation.

We then used the simple slope test to analyze the conditional effect of technology on the impact of loneliness on the perception of AFEs to further understand the impact. As evidenced in Figure 4, the link between loneliness and AFEs was not as obvious when technology was low (β simple: mean −1, SD −0.0819, t_1=−1.5394; *P*=.13). Conversely, when technology was high, the impact of moderation was apparent more clearly (β simple: mean 1, SD −0.4983, t_{90}=−7.2636; *P*<.001). Thus, we can conclude that technology moderates the impact of loneliness on the perception of AFEs when the technology experience is high, confirming H4.

Figure 3. Moderating effect of technology experience in the relationship between loneliness and age-friendliness of environments after controlling for life-space mobility.

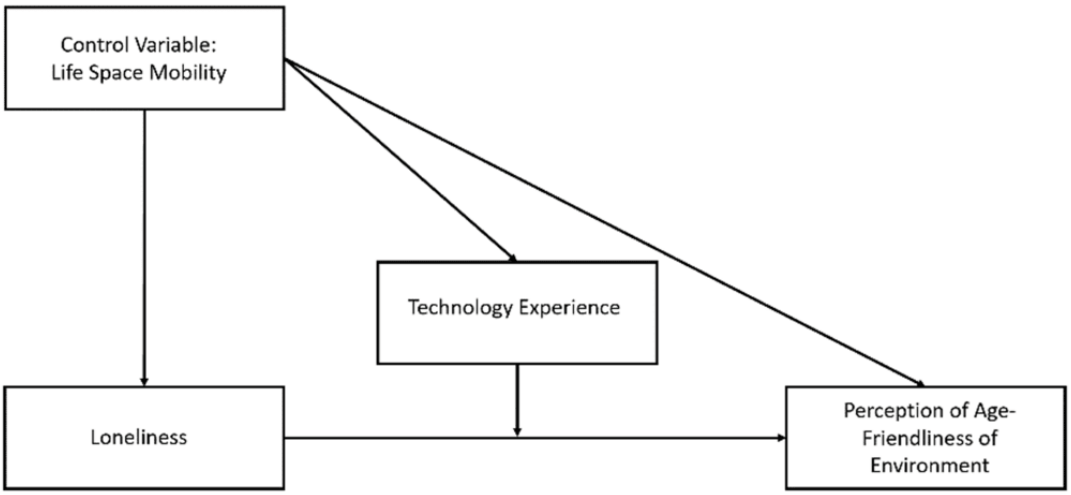


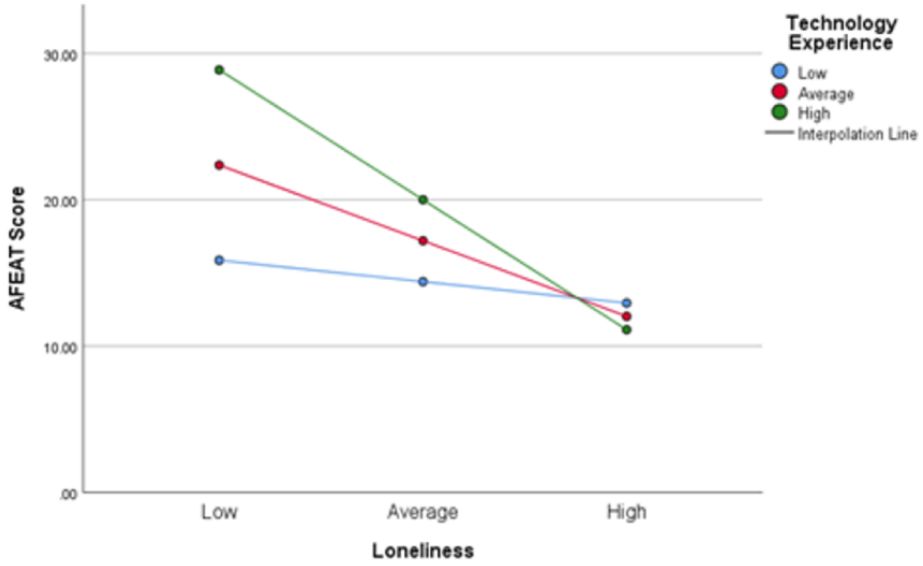
Table 4. The relationship between loneliness and technology experience and the impact of life-space mobility.

	Overall fit indicators			Significance of standardized coefficient	
	R	R^2	*F* test (*df*)	β	*t* test
Variables	0.82	0.68	45.8722 (1,90)		
Loneliness (ZL1)	—[a]	—	—	−0.5335	−6.2794[b]
Technology experience (ZTE1)	—	—	34.78 (1,90)	0.2888	3.5746[b]
ZL1*ZTE1	—	—	—	−0.3829	−5.1518[b]
Life-space mobility (ZLS1)	—	—	—	0.1177	1.665

[a]Not applicable.

[b]*P*<.001.

Figure 4. Moderating effect of technology experience in the relationship between loneliness and age-friendliness of environments after controlling for life-space mobility (simple slope test).



Discussion

Principal Findings

This study found higher levels of loneliness during the height of the COVID-19 pandemic among the older adult sample compared with the commonly available prepandemic data. For instance, Victor and Bowling [59] reported an average loneliness level of 30%, whereas Hawkley and Cacioppo [60] found the prevalence of loneliness in older adults to be approximately 25%, compared with the 44% found in this study. Numerous studies have found an increase in loneliness during the pandemic [61-63].

Higher loneliness was found to be correlated with a more negative perception of AFE, confirming H2 and in line with what has been theorized by the prevailing literature [18-20,22]. Shortfalls in emotional and social fulfillment have been highlighted as a predictor of loneliness, and the lack of opportunities for older adults to socialize, access important key services, and ability to engage with the community may have worsened perceptions. When older adults perceive that their desired quantity or quality of social engagements is met, they are less likely to experience loneliness and may also have favorable attitudes toward AFEs. Previous research has suggested that this relationship is likely to be complex [19] and bidirectional. AFE perception can be an indicator of a mechanism for aging-in-place, preserving older adults' physical setting of choice, providing a sense of community attachment, and allowing them to engage with a developed social network, with associated familiarity and better mental well-being.

This study found that greater use of technology clearly moderated the relationship between loneliness and the perception of AFEs. It also moderated the LSM-AFE relationship. Furthermore, the ability to use technology successfully to adapt to challenging experiences during lockdown emerged as a potential buffer against the impact of loneliness on AFE perception. Other studies that examined technology use during the pandemic revealed an increase, indicating adaptation to LSM restrictions via alternative pathways [37,64]. An increase in LSM restriction was linked to a positive perception of AFEs, rejecting H1, which at first seemed counterintuitive. However, the moderation effect of technology, such that this relationship existed only when technology use was high, explained the unexpected direction of the relationship. Older adults were overcoming physical restrictions barriers, where the ability to replace previously in-person activities with those online may have impacted perceptions. Other studies that examined AFEs also noted an increase in positive perception of AFEs as the pandemic progressed, but these could also have been linked to the easing of restrictions [56]. Overall, technology use may have improved the negative impacts of loneliness on the perception of AFEs helping with older adult mental health.

In testing H4, we determined that when technology experience was high, it had a moderation effect on the loneliness-AFE relationship. Ng et al [16] attempted to explain the relationship between internet use and loneliness through a moderation-mediation mechanism between internet use, perception of AFE, and loneliness, and depression. Their findings were consistent with those of Park et al [24] and Domènech-Abella et al [65], who confirmed the moderating effect of the internet on the age-friendliness-loneliness-depression mechanism, which may also explain our results. For example, Booth et al [66] discovered a partial mediation effect of feelings of helplessness, social isolation, depression, and distrust between psychological distress and perceived AFE (especially concerning security). Social isolation is a well-established predictor of loneliness and depression, and loneliness has also been linked in the literature. Taken together, these hint at a causal route between loneliness and perception of AFEs. Increased anxiety can lead to increased loneliness and mental health issues, potentially leading to a reduced fit between the older adult and the environment [67,68].

Despite studies highlighting the potential moderating role of technology between loneliness and AFE perception, previous results were always unclear about a direct link between these 3 variables. Our study found a clear moderating role for technology in the relationship between loneliness and AFEs, which is a notable finding.

When using the simple slope test, where technology use was low, we found that LSM negatively affected perceived AFEs when technology use was low. This explained our initial counterintuitive result and supported earlier findings [24]. To substitute or overcome confinement, older adults may have developed alternative routes of access to AFE domains, such as social, civic engagement, and access to services through technology.

Access to the internet may have helped reduce older people's boredom, and there have been examples from recent studies where older adults were able to access informal help networks, reading, or online game groups, as well as participate in community-based activities like attending virtual church gatherings [69,70]. Alternative social and emotional outlets to combat loneliness through access via the internet, to previously in-person services (eg, primary care, counseling) may have contributed to network socialization, allowing older adults to continue feeling like they are a part of their environment and reducing loneliness.

The implications of our study can be applied to situations outside of the pandemic context. Studies have highlighted that older adults prefer to live in their own homes and interact with their local community, where they have developed relationships over time and do so as long as possible [71,72]. A negative perception of AFEs can be viewed as a barometer indicating poor person-environment fit [72], associated with poorer mental health outcomes [25,71]. Therefore, technology could be a solution for those who ordinarily have limited LSM, are at risk for social isolation, and have a negative perception of AFEs.

Our study confirmed the link between LSM, loneliness, and perception of AFEs, as well as the moderating role technology played during the COVID-19 pandemic, advancing the findings of previous studies in this rapidly evolving body of literature. We also strengthened the argument that the WHO's Global Age-Friendly Cities framework would benefit older adults more by including technology use as an additional ninth domain.

Study Limitations

This study had several limitations. For example, the sampling could have been predisposed to participants literate in digital resources and more socially connected. Generally, such participants may have experienced less loneliness [73]. Furthermore, a cross-sectional design cannot establish causality [74]. Although the sample size was small, the statistical power, effect size, precision, type, complexity of analysis, study population variability, and homogeneity overcame this shortcoming. The results are an important contribution to the discourse on the role being played by AFEs in the mental well-being of older adults and the role of technology.

Although it is unlikely that participants will experience the same level of LSM restrictions after the pandemic, this allowed us to examine the studied measures in a normally inaccessible environment. However, we cannot conclude with certainty that the pandemic caused the observations because we did not have a prepandemic assessment for the same participants. Nevertheless, other prepandemic studies supported our hypothesis of mobility restrictions' impact on loneliness.

Conclusion

Despite the limitations, the results suggest that the vulnerability of older adults during the pandemic and their exposure to loneliness and negative perceptions of AFEs increased, and technology played an important role in moderating these influences. Given the advancements in technology, the WHO's 8 domains of AFEs may be obsolete with their narrow implicit recognition of technology. Our study demonstrates the significance of an explicit recognition of technology in the evaluation of AFEs as an integral component of all aspects of older adults' daily lives. Future studies may also wish to look further into the impact of demographics and differences between genders.

Community and mental health service access could be improved by providing online access. Older adults need cheap access to internet infrastructure, and health community centers should provide technology training, attend online meetings, or use mental health applications.

Researchers in the field of loneliness in older adults are encouraged to use our results to inform initiatives to reduce the mental health risks for older adults in vulnerable crises, such as the pandemic and civil insurrections [75,76]. Appropriate consideration of these factors will aid decision makers in developing robust and effective strategies during times of crisis, as well as in assisting an aging population with aging-in-place.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE Checklist.
[DOCX File , 35 KB - aging_v8i1e67242_app1.docx]

References

1. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2007-10-05]
2. Phillipson C, Baars J. Social theory and social ageing. Ageing Soc 2007;22:68-84. [doi: 10.4135/9781446278918.n4]
3. Phillipson C. Developing age-friendly urban communities: critical issues for public policy. Public Policy Aging Rep 2014 Dec 27;25(1):4-8. [doi: 10.1093/ppar/pru052]
4. Buffel T, Phillipson C, scharf T. Ageing in urban environments: developing 'age-friendly' cities. Crit Soc Policy 2012 May 22;32(4):597-617. [doi: 10.1177/0261018311430457]
5. Phillipson C. Developing Age-Friendly Communities: New Approaches to Growing Old in Urban Environments. Berlin, Heidelberg: Springer; 2010.
6. Cerin E, Nathan A, van Cauwenberg J, Barnett DW, Barnett A, Council on EnvironmentPhysical Activity (CEPA) – Older Adults working group. The neighbourhood physical environment and active travel in older adults: A systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017;14(1):15 [FREE Full text] [doi: 10.1186/s12966-017-0471-5] [Medline: 28166790]
7. Young AF, Russell A, Powers JR. The sense of belonging to a neighbourhood: can it be measured and is it related to health and well being in older women? Soc Sci Med 2004;59(12):2627-2637. [doi: 10.1016/j.socscimed.2004.05.001] [Medline: 15474215]
8. Garner IW, Holland CA. Age-friendliness of living environments from the older person's viewpoint: development of the age-friendly environment assessment tool. Age Ageing 2020;49(2):193-198. [doi: 10.1093/ageing/afz146] [Medline: 31790132]
9. Balki E, Hayes N, Holland C. The impact of social isolation, loneliness, and technology use during the COVID-19 pandemic on health-related quality of life: observational cross-sectional study. J Med Internet Res 2022;24(10):e41536 [FREE Full text] [doi: 10.2196/41536] [Medline: 36260401]
10. Mahmood A, Seetharaman K, Jenkins HT, Chaudhury H. Contextualizing innovative housing models and services within the age-friendly communities framework. Gerontologist 2022;62(1):66-74. [doi: 10.1093/geront/gnab115] [Medline: 34355769]
11. Hsu HC. Associations of city-level active aging and age friendliness with well-being among older adults aged 55 and over in Taiwan. Int J Environ Res Public Health 2020;17(12):4304 [FREE Full text] [doi: 10.3390/ijerph17124304] [Medline: 32560170]
12. Berg-Weger M, Morley JE. Editorial: Loneliness and social isolation in older adults during the COVID-19 pandemic: implications for gerontological social work. J Nutr Health Aging 2020;24(5):456-458 [FREE Full text] [doi: 10.1007/s12603-020-1366-8] [Medline: 32346678]
13. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. International Psychogeriatrics 2020;32(10):1217-1220. [doi: 10.1017/s1041610220000988]
14. Au A, Lai DWL, Yip H, Chan S, Lai S, Chaudhury H, et al. Sense of community mediating between age-friendly characteristics and life satisfaction of community-dwelling older adults. Front Psychol 2020;11:86 [FREE Full text] [doi: 10.3389/fpsyg.2020.00086] [Medline: 32194465]
15. Aboderin I, Kano M, Owii HA. Toward "Age-Friendly Slums"? Health challenges of older slum dwellers in nairobi and the applicability of the age-friendly city approach. Int J Environ Res Public Health 2017;14(10):1259 [FREE Full text] [doi: 10.3390/ijerph14101259] [Medline: 29053576]
16. Ng SI, Lim XJ, Hsu HC, Chou CC. Age-friendliness of city, loneliness and depression moderated by internet use during the COVID-19 pandemic. Health Promot Int 2023;38(3) [FREE Full text] [doi: 10.1093/heapro/daac040] [Medline: 35437585]
17. Perlman D, Peplau LA. In: Gilmour R, Duck S, editors. Toward a social psychology of loneliness. In Personal Relationships: 3. Relationships in Disorder. Cambridge, Massachusetts, USA: Academic Press; 1981:56.
18. Dinkins CS. Seeing oneself in the face of the other: the value and challenge of human connectedness for older adults. J Psychosoc Nurs Ment Health Serv 2017;55(7):13-17. [doi: 10.3928/02793695-20170619-03] [Medline: 28671237]
19. Kearns A, Whitley E, Tannahill C, Ellaway A. 'Lonesome Town'? Is loneliness associated with the residential environment, including housing and neighbourhood factors? J Community Psychol 2015;43(7):849-867 [FREE Full text] [doi: 10.1002/jcop.21711] [Medline: 26740728]
20. Lee MS, Park J. Television shopping at home to alleviate loneliness among older consumers. amj 2017;18(4):139. [doi: 10.15830/amj.2017.18.4.139]
21. Gibney S, Zhang M, Brennan C. Age-friendly environments and psychosocial wellbeing: a study of older urban residents in Ireland. Aging Ment Health 2020;24(12):2022-2033. [doi: 10.1080/13607863.2019.1652246] [Medline: 31402677]
22. Rantakokko M, Iwarsson S, Vahaluoto S, Portegijs E, Viljanen A, Rantanen T. Perceived environmental barriers to outdoor mobility and feelings of loneliness among community-dwelling older people. J Gerontol A Biol Sci Med Sci 2014;69(12):1562-1568. [doi: 10.1093/gerona/glu069] [Medline: 24864307]
23. Stafford PB. The Global Age-Friendly Community Movement: A Critical Appraisal, 5. New York: Berghahn Books; 2018.
24. Park S, Kim M, Chung S. Age-friendly environments and depressive symptoms among Korean adults: the mediating effects of loneliness. Aging Ment Health 2021;25(6):1060-1070. [doi: 10.1080/13607863.2020.1755827] [Medline: 32321293]

25. Gaugler JE, Duval S, Anderson KA, Kane RL. Predicting nursing home admission in the U.S: a meta-analysis. BMC Geriatr 2007;7(1):13 [FREE Full text] [doi: 10.1186/1471-2318-7-13] [Medline: 17578574]
26. Stalvey BT, Owsley C, Sloane ME, Ball K. The life space questionnaire: a measure of the extent of mobility of older adults. J Appl Gerontol 1999;18(4):460-478. [doi: 10.1177/073346489901800404]
27. Xue QL, Fried LP, Glass TA, Laffan A, Chaves PHM. Life-space constriction, development of frailty, and the competing risk of mortality: the Women's Health And Aging Study I. Am J Epidemiol 2008;167(2):240-248. [doi: 10.1093/aje/kwm270] [Medline: 17906296]
28. Sachs J. Sustainable Development Report 2020: the Sustainable Development Goals and Covid-19 Includes the SDG Index and Dashboards. Cambridge: Cambridge University Press; 2021.
29. Kotwal AA, Holt-Lunstad J, Newmark RL, Cenzer I, Smith AK, Covinsky KE, et al. Social isolation and loneliness among san francisco bay area older adults during the COVID-19 shelter-in-place orders. J Am Geriatr Soc 2021;69(1):20-29 [FREE Full text] [doi: 10.1111/jgs.16865] [Medline: 32965024]
30. Emlet CA, Moceri JT. The importance of social connectedness in building age-friendly communities. J Aging Res 2012;2012:1-9 [FREE Full text] [doi: 10.1155/2012/173247] [Medline: 22162807]
31. Lui CW, Everingham JA, Warburton J, Cuthill M, Bartlett H. What makes a community age-friendly: A review of international literature. Australas J Ageing 2009;28(3):116-121. [doi: 10.1111/j.1741-6612.2009.00355.x] [Medline: 19845650]
32. Peel C, Sawyer Baker P, Roth DL, Brown CJ, Brodner EV, Allman RM. Assessing mobility in older adults: the UAB study of aging life-space assessment. Phys Ther 2005;85(10):1008-1119. [Medline: 16180950]
33. Theis KA, Furner SE. Shut-In? Impact of chronic conditions on community participation restriction among older adults. J Aging Res 2011;2011 [FREE Full text] [doi: 10.4061/2011/759158] [Medline: 21837277]
34. Rantakokko M, Iwarsson S, Kauppinen M, Leinonen R, Heikkinen E, Rantanen T. Quality of life and barriers in the urban outdoor environment in old age. J Am Geriatr Soc 2010;58(11):2154-2159. [doi: 10.1111/j.1532-5415.2010.03143.x] [Medline: 21054297]
35. Coyle CE, Dugan E. Social isolation, loneliness and health among older adults. J Aging Health 2012;24(8):1346-1363. [doi: 10.1177/0898264312460275]
36. Groarke JM, Berry E, Graham-Wisener L, McKenna-Plumley PE, McGlinchey E, Armour C. Loneliness in the UK during the COVID-19 pandemic: Cross-sectional results from the COVID-19 psychological wellbeing study. PLoS One 2020;15(9):e0239698 [FREE Full text] [doi: 10.1371/journal.pone.0239698] [Medline: 32970764]
37. Moore RC, Hancock JT. Older adults, social technologies, and the coronavirus pandemic: challenges, strengths, and strategies for support. Soc Media Soc 2020 Aug 11;6(3). [doi: 10.1177/2056305120948162]
38. Shah SGS, Nogueras D, van Woerden HC, Kiparoglou V. Evaluation of the effectiveness of digital technology interventions to reduce loneliness in older adults: systematic review and meta-analysis. J Med Internet Res 2021;23(6):e24712 [FREE Full text] [doi: 10.2196/24712] [Medline: 34085942]
39. Dowson B, Atkinson R, Barnes J, Barone C, Cutts N, Donnebaum E, et al. Digital approaches to music-making for people with dementia in response to the COVID-19 pandemic: current practice and recommendations. Front Psychol 2021;12:625258 [FREE Full text] [doi: 10.3389/fpsyg.2021.625258] [Medline: 33967893]
40. Feldman PH, Oberlink MR. Developing community indicators to promote the health and well-being of older people. Fam Community Health 2003;26(4):268-274. [doi: 10.1097/00003727-200310000-00004] [Medline: 14528133]
41. Lawton MP. Environment and other determinants of well-being in older people. Gerontologist 1983;23(4):349-357. [doi: 10.1093/geront/23.4.349] [Medline: 6352420]
42. Gerst-Emerson K, Jayawardhana J. Loneliness as a public health issue: the impact of loneliness on health care utilization among older adults. Am J Public Health 2015;105(5):1013-1019. [doi: 10.2105/ajph.2014.302427]
43. Luciano A, Pascale F, Polverino F, Pooley A. Measuring age-friendly housing: a framework. Sustainability 2020;12(3):848. [doi: 10.3390/su12030848]
44. Marston HR, van Hoof J. "Who Doesn't think about technology when designing urban environments for Older People?" a case study approach to a proposed extension of the WHO's age-friendly cities model. Int J Environ Res Public Health 2019;16(19):3525 [FREE Full text] [doi: 10.3390/ijerph16193525] [Medline: 31547204]
45. Sánchez-González D, Rojo-Pérez F, Rodríguez-Rodríguez V, Fernández-Mayoralas G. Environmental and psychosocial interventions in age-friendly communities and active ageing: A systematic review. Int J Environ Res Public Health 2020;17(22):8305 [FREE Full text] [doi: 10.3390/ijerph17228305] [Medline: 33182710]
46. Balki E, Hayes N, Holland C. Effectiveness of technology interventions in addressing social isolation, connectedness, and loneliness in older adults: Systematic umbrella review. JMIR Aging 2022;5(4):e40125 [FREE Full text] [doi: 10.2196/40125] [Medline: 36279155]
47. Belam G. Yoga as an intervention for older peoples mental health: a literature review. WWOP 2020;24(3):159-169. [doi: 10.1108/wwop-05-2020-0017]
48. Pedell S, Borda A, Keirnan A, Aimers N. Combining the Digital, Social and Physical Layer to Create Age-Friendly Cities and Communities. Int J Environ Res Public Health 2021;18(1):325 [FREE Full text] [doi: 10.3390/ijerph18010325] [Medline: 33466259]

49. Reuter A, Liddle J, Scharf T. Digitalising the age-friendly city: Insights from participatory action research. Int J Environ Res Public Health 2020;17(21):8281 [FREE Full text] [doi: 10.3390/ijerph17218281] [Medline: 33182453]
50. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
51. Allen M. The SAGE Encyclopedia of Communication Research Methods. Washington DC: SAGE; 2017.
52. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
53. Czaja SJ, Charness N, Fisk AD, Hertzog C, Nair SN, Rogers WA, et al. Factors predicting the use of technology: findings from the center for research and education on aging and technology enhancement (CREATE). Psychol Aging 2006;21(2):333-352 [FREE Full text] [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
54. Hayes AF. Partial, conditional, and moderated mediation: quantification, inference, and interpretation. Commun Monogr 2018;85:4-40.
55. Mullins LC. Encyclopedia of Gerontology, 2nd ed. Berlin, Heidelberg,: Springer; 2007.
56. Garner IW, Varey S, Navarro-Pardo E, Marr C, Holland CA. An observational cohort study of longitudinal impacts on frailty and well-being of COVID-19 lockdowns in older adults in England and spain. Health Soc Care Community 2022;30(5):e2905-e2916. [doi: 10.1111/hsc.13735] [Medline: 35089638]
57. Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. J Counsel Psychol 2004 Jan;51(1):115-134. [doi: 10.1037/0022-0167.51.1.115]
58. Aiken LS, West SG. Multiple Regression: Testing and Interpreting Interactions. Washington DC: SAGE; 1991.
59. Victor CR, Bowling A. A longitudinal analysis of loneliness among older people in Great Britain. J Psychol 2012;146(3):313-331. [doi: 10.1080/00223980.2011.609572] [Medline: 22574423]
60. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010;40(2):218-227 [FREE Full text] [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
61. Bu F, Steptoe A, Fancourt D. Loneliness during a strict lockdown: Trajectories and predictors during the COVID-19 pandemic in 38,217 United Kingdom adults. Soc Sci Med 2020;265:113521 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113521] [Medline: 33257177]
62. Elran-Barak R, Mozeikov M. One month into the reinforcement of social distancing due to the COVID-19 outbreak: subjective health, health Behaviors, and Loneliness among people with chronic medical conditions. Int J Environ Res Public Health 2020;17(15):5403 [FREE Full text] [doi: 10.3390/ijerph17155403] [Medline: 32727103]
63. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being. J Gerontol B Psychol Sci Soc Sci 2021;76(2):e53-e58 [FREE Full text] [doi: 10.1093/geronb/gbaa110] [Medline: 32778899]
64. Balki E, Hayes N, Holland C. Loneliness and older adults: psychological resilience and technology use during the COVID-19 pandemic-a cross sectional study. Front Aging 2023;4:1184386 [FREE Full text] [doi: 10.3389/fragi.2023.1184386] [Medline: 37434741]
65. Domènech-Abella J, Switsers L, Mundó J, Dierckx E, Dury S, De Donder L. The association between perceived social and physical environment and mental health among older adults: mediating effects of loneliness. Aging Ment Health 2021;25(5):962-968. [doi: 10.1080/13607863.2020.1727853] [Medline: 32067471]
66. Booth J, Ayers SL, Marsiglia FF. Perceived neighborhood safety and psychological distress: exploring protective factors. The Journal of Sociology & Social Welfare 2012;39(4):137. [doi: 10.15453/0191-5096.3703]
67. Balki E, Holland C, Hayes N. A mixed methods exploration of older adult use and acceptance of digital communication technology for social connectedness during the Covid-19 pandemic,. J Med Internet Res 2023;25:e41535 [FREE Full text] [doi: 10.2196/41535] [Medline: 37531187]
68. Balki E, Hayes N, Holland C. The indirect impact of educational attainment as a distal resource for older adults on loneliness, social isolation, psychological resilience, and technology use during the COVID-19 pandemic: cross-sectional quantitative study. JMIR Aging 2023;6:e47729 [FREE Full text] [doi: 10.2196/47729] [Medline: 37999938]
69. Cohen-Mansfield J, Muff A, Meschiany G, Lev-Ari S. Adequacy of web-based activities as a substitute for in-person activities for older persons during the COVID-19 pandemic: survey study. J Med Internet Res 2021 Jan 22;23(1):e25848 [FREE Full text] [doi: 10.2196/25848] [Medline: 33439851]
70. Gibney S, Moore T, Shannon S. Loneliness in later life: a cross-sectional survey analysis of place-based factors in Ireland. QAOA 2019;20(2):80-96. [doi: 10.1108/qaoa-04-2018-0015]
71. Greenfield EA. Support from neighbors and aging in place: can NORC programs make a difference? Gerontologist 2016;56(4):651-659. [doi: 10.1093/geront/gnu162] [Medline: 26035886]
72. Park H, Park S, Shephard RJ, Aoyagi Y. Yearlong physical activity and sarcopenia in older adults: the nakanojo study. Eur J Appl Physiol 2010;109(5):953-961. [doi: 10.1007/s00421-010-1424-8] [Medline: 20336310]
73. Chronic PSO. Interventions to combat loneliness for older adults. A review of reviews 2018;964((Suppl. 1)). [doi: 10.1093/geroni/igy031.3571]

74. Sygit-Kowalkowska E, Szrajda J, Weber-Rajek M, Porażyński K, Ziółkowski M. Resilience as a predicator of mental health of incarcerated women. Psychiatr Pol 2016;51(3):549-560. [doi: 10.12740/pp/onlinefirst/62617]
75. Abdulkerim S, Albari M, MacKenzie A, Thurston A. Title: report on a survey of mobility disability among Syrian refugees living in Turkey. Int J Educational Res Open 2021;2:100059. [doi: 10.1016/j.ijedro.2021.100059]
76. Fund KS. Access to Health Care and Minority Ethnic Groups. (King?s Fund). 2006. URL: https://assets.kingsfund.org.uk/f/256914/x/9e3f53c8f1/access_health_care_minority_ethnic_groups_2006.pdf [accessed 2025-04-03]

Abbreviations

AFE: age-friendly environment
AFEAT: Age-Friendly Environment Assessment Tool
H: hypothesis
LSM: life-space mobility
QOL: quality of life
SAGE: Scientific Advisory Group for Emergencies
STROBE: Strengthening Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by A Quialheiro; submitted 06.10.24; peer-reviewed by C Yang, C Franceschetti; comments to author 29.01.25; revised version received 25.02.25; accepted 02.04.25; published 06.05.25.

Please cite as:
Balki E, Hayes N, Holland C
Factors Influencing Older Adults' Perception of the Age-Friendliness of Their Environment and the Impact of Loneliness, Technology Use, and Mobility: Quantitative Analysis
JMIR Aging 2025;8:e67242
URL: https://aging.jmir.org/2025/1/e67242
doi:10.2196/67242
PMID:

Digital Literacy and Its Association With Subjective Health Status and Healthy Lifestyle Behaviors Among Korean Older Adults: Cross-Sectional Study

Soon Young Lee[1*], MPH; Yejin Kim[2*], MPH; Bomgyeol Kim[3], PhD; Sang Gyu Lee[4], PhD; Suk-Yong Jang[5], PhD; Tae Hyun Kim[6], PhD

[1]Department of Strategic Planning Part, Kangbuk Samsung Hospital, Seoul, Republic of Korea

[2]Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

[3]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[4]Department of Preventive Medicine, College of Medicine, Yonsei University, Seoul, Republic of Korea

[5]Department of Biohealth Policy Analysis, Graduate School of Transdisciplinary Health Sciences, Yonsei University, Seoul, Republic of Korea

[6]Department of Biohealth Industry Management, Graduate School of Transdisciplinary Health Sciences, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Tae Hyun Kim, PhD
Department of Biohealth Industry Management, Graduate School of Transdisciplinary Health Sciences, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: With an aging population driven by advances in medical technology, digital literacy has become essential for improving the quality of life of older adults, enhancing access to health information, and promoting healthy lifestyles. Furthermore, the COVID-19 pandemic may have influenced the subjective health perceptions and healthy lifestyle behaviors of older adults. However, there is limited research exploring the relationship between digital literacy, subjective health perceptions, and healthy lifestyle behaviors in Korea.

Objective: This study aimed to investigate digital literacy's impact on Korean older adults' subjective health status and healthy lifestyle behaviors.

Methods: Data of 8664 respondents (aged 65 years and older) from the 2020 National Survey of the Older Koreans were analyzed. Digital literacy was measured based on the use of IT devices (ITDs), difficulty using online information, and inconvenience of ITDs. Statistical analyses, such as the Rao-Scott chi-square test, Wilcoxon rank sum test, and multiple regression analysis, were conducted.

Results: Respondents with above-average ITD use (adjusted odds ratio [aOR] 1.73, 95% CI 1.50 - 1.99) and less difficulty using online information (aOR 1.41, 95% CI 1.24 - 1.61) had higher odds of perceiving themselves as healthy. Conversely, high difficulty using ITDs was associated with lower odds of respondents perceiving themselves as healthy (aOR 0.84, 95% CI 0.82 - 0.87). Furthermore, high ITD use predicted engagement in healthy lifestyle behaviors (aOR 1.51, 95% CI 1.33 - 1.72), whereas high difficulty using ITDs predicted lower odds of engagement (aOR 0.94, 95% CI 0.92 - 0.97). In contrast, there was no difference in the odds of engaging in healthy lifestyle behaviors regardless of difficulty using online information (aOR 1.03, 95% CI 0.92 - 1.15).

Conclusions: This study underscores the significant association between digital literacy and improved health outcomes among older adults. Promotion of digital literacy and relevant policies is essential to help older adults effectively obtain health information online, thereby improving their quality of life and overall health.

(JMIR Aging 2025;8:e64974) doi:10.2196/64974

KEYWORDS

digital literacy; healthy lifestyle behaviors; older adults; subjective health status; quality of life

Introduction

Advances in medical technology have increased life expectancy, thereby contributing to the growth of an aging population worldwide [1]. South Korea became an aging society in 2017, and it is predicted to transition to a super-aged society by 2025 [2]. Old age is characterized by a general decline in physical and mental functioning, leading to a range of health issues, including senile and chronic degenerative diseases and impaired daily functioning [3]. Prevention, treatment, and management of diseases are necessary to improve of older adults' health care and quality of life. In addition, health information and proactive responses for health care are important [4]. Particularly, the first generation of Korean "baby boomers" (born between 1955 and 1972), specifically those born in 1955, has been incorporated into the older adult population since 2020. Thus, this is a crucial time for discussion on older adults' health care.

Generally, older adults obtain health information from health care providers, friends and relatives, television, newspapers, radio, and other mass media to maintain their health [5]. However, rapid technological developments have increased the opportunities for acquiring health information and news on the internet, leading to digital literacy's increased importance [6-8]. Digital technology initially spread among young people, men, and highly educated people and is gradually spreading to older adults, women, and less educated people. The aging trend and the rapid development of new media technology suggest that older adults will have to use digital technology more in the future [9]. Digital literacy comprises the ability to understand the information generated on a digital device and implement it accordingly [10]. In an information society, digital literacy's importance is expanding to all areas of life as a key factor in improving quality of life [11,12].

The manner in which older adults perceive their health informs how they relate to society. In addition, their subjective well-being is intimately linked to their perceived health status [13]. Meanwhile, health concerns can be a primary motivation for searching health information online [14]. Specifically, the use of the internet promotes the health of older adults by improving access to health information, encouraging healthy lifestyles, and enhancing social interaction [15]. While older adults recognize the effectiveness and usefulness of using health information through digital media such as the internet for health care, they experience challenges when actually attempting to obtain health information online [16,17]. According to a 2022 report on the digital divide, vulnerable groups such as people with disabilities, low-income individuals, farmers, and older adults have markedly lower digital use capabilities, with the latter experiencing the highest level of information vulnerability [18].

The growing importance of digital literacy for older adults has prompted a variety of studies. However, most Korean research is survey-based or restricted to specific geographic regions or disease populations, focusing primarily on health knowledge, mental health aspects such as depression and self-efficacy, or quality of life [19-22]. While subjective health perceptions and healthy lifestyle behaviors of older adults may have been influenced by COVID-19 pandemic, these studies were conducted before the pandemic [23-25]. Furthermore, the rapid progress of aging and the advancement of IT, along with gaps in the acquisition of information, unchecked spread of health information, and dissemination of unverified information, have created growing concerns among the older adult population vis-à-vis the lack of reliable information [26-28].

This study seeks to examine 2 key hypotheses: (1) older adults with higher levels of digital literacy are more likely to perceive themselves as healthy and (2) higher levels of digital literacy among older adults are associated with a greater likelihood of engaging in health-promoting behaviors. Using a nationally representative sample, we aim to investigate the association between different aspects of digital literacy and older adults' subjective health status and healthy lifestyle behaviors.

Methods

Data Source and Study Population

This cross-sectional study used data from the 2020 National Survey of the Older Koreans (NSOK), conducted by the Ministry of Health and Welfare and the Korea Institute for Health and Social Affairs (KIHASA). The NSOK is a survey conducted every 3 years since 2008, in accordance with Article 5 of the Welfare of Senior Citizens Act. It systematically collects data from more than 10,000 older adults aged 65 years and older who are living in ordinary residences across 17 cities and provinces nationwide. Trained surveyors carry out the survey in 969 districts selected based on regional representation. They survey all households with older residents, considering factors such as area of residence, age, gender, education, and marital status [29]. This survey aimed to identify the living conditions, characteristics, and needs of older adults to develop welfare policies to improve their quality of life. In addition, it sought to develop policies in response to Korea's aging society by examining changes in the characteristics of older adults through the accumulation of time-series data [29].

NSOK calculates weights in 3 steps: using design weights, nonresponse-adjusted weights, and poststratification weights. First, the design weights are directly derived from the sampling design, using a 2-stage cluster sampling method. This method involves probability proportional to size (PPS) sampling, where sample areas are selected based on the number of households in each area, followed by the selection of households within each sampled area. Nonresponse adjustment is conducted by upweighting responding households using the reciprocal of the response rate. Final weights are calculated by applying a raking ratio method, and respondent weights, adjusted for design and nonresponse weights, are poststratified to match the distribution of older adults in the population, thereby enhancing the accuracy of estimates. The 2020 NSOK used data from the 2019 Census, the most recent available at the time of survey completion, to calculate final weights by applying population data for older adults by gender, age group (65 - 74 years or 75 years and older), and household size (1-person household, 2+ persons) across the 17 metropolitan areas [21].

The study excluded 1256 of the 10,097 respondents who answered "don't know" or "not applicable" to the digital literacy survey. In addition, if the respondent has a physical or mental illness, the cohabitant or a close representative of the study subject responds on their behalf. In this case, only the objective information of the older adults was surveyed, and the main focus of this study was not surveyed [28]. Therefore, the results were generalized to older Koreans by excluding the 177 respondents who responded on behalf of their cohabitant. These exclusions ensured the objectivity of the health indicator responses, as the survey focused on the subjective judgment of the original respondents [29]. Finally, the study included 8664 participants.

Ethical Considerations

The Institutional Review Board (IRB) of the Yonsei Medical Center approved a consent exemption for this study (No. 4-2023-0361). The 2020 Elderly Survey used in this study was anonymized by the data provider, KIHASA, and the tools and processes used in the survey were approved by the IRB of KIHASA (No. 2020 - 36).

Measurements

Subjective Health Status and Healthy Lifestyle Behaviors

The dependent variables of this study included subjective health status and healthy lifestyle behaviors. Subjective health status has been widely used in various studies as a valid proxy for actual health status, given its reliable predictive value for health outcomes [30-32]. It was categorized into good (very healthy and healthy) and poor (fair, poor, and very poor). Healthy lifestyle behaviors were defined as smoking cessation, decreased alcohol consumption, and exercise [33]. Smoking cessation was defined as having quit smoking, decreased alcohol consumption was defined as not drinking at all in the previous year or drinking less than twice a week, and exercise was defined as answering "yes" to the question, "Do you usually exercise for at least 10 minutes on a consistent basis?"

Digital Literacy

This study's main variable was digital literacy, which was divided into 3 subcategories. It was scored based on the degree of use of IT devices (ITDs), difficulty using information provided online, and inconvenience of ITD usage.

For the item on the degree of ITD usage, 1 point was assigned to each of the activities in Textbox 1, which involved using a PC, mobile phone, or tablet. The sample was divided into 2 groups according to the respondent's score, namely above average (3 points or more) and below average (3 points or less).

For the item on the difficulty of using information provided online, the degree of difficulty in using information or services required for daily life via the internet was categorized into "difficult" and "not difficult." The inconvenience of ITD usage was assessed considering tasks such as making online reservations, conducting purchases using machines such as kiosks and automated teller machines, and shopping at stores that only accept cards. For each task, the level of inconvenience was scored on a scale from 0 ("not at all inconvenient") to 2 ("inconvenient"). Finally, the overall level of inconvenience of using ITDs was rated on a scale from 0 ("not at all inconvenient") to 8 ("inconvenient"). A Cronbach α of 0.88 was identified for ITD usage, and 0.80 for inconvenience of using ITDs.

Textbox 1. Activities that involved using a PC, mobile phone, or tablet.

- Receiving messages.
- Sending messages.
- Searching for and viewing information.
- Taking photos and videos.
- Listening to music.
- Playing games.
- Watching videos.
- Using social media.
- Using e-commerce platforms.
- Using online financial transactions.
- Searching for and installing applications.
- Other activities.

Covariates

The covariates included sociodemographic characteristics and health-related factors. Sociodemographic characteristics included gender, age, residence, educational level, household types, and household income. Age was classified into five 5-year intervals (65 - 69, 70 - 74, 75 - 79, 80 - 84, and ≥85 years old). Residence was classified into "rural (eup/myeon)" and "urban (dong)." Educational status was classified into "elementary school or less," "middle and high school," and "university and above." Household type was classified into "single," "married couple," and "cohabitant (nonspouse)." Household income encompassed all household members' earned, business, property, and transfer incomes. This comprised the annualized gross household income for the most recent year (July 1, 2019, to

June 30, 2020), converted to a monthly basis and divided into quartiles.

Health-related factors included the number of chronic diseases respondents had. Chronic diseases included conditions that lasted for more than 3 months and had been diagnosed by a doctor. These were categorized as "none," "1‐2 chronic diseases," and "3 or more chronic diseases," encompassing 32 conditions.

Statistical Analysis

A frequency analysis was conducted to identify the participants' general characteristics. The Rao-Scott chi-square test and Wilcoxon rank sum test were conducted to evaluate and compare these characteristics. Multiple logistic regression analysis examined the associations of digital literacy with subjective health status and healthy lifestyle behaviors. In addition, a subgroup analysis was conducted to determine whether gender led to differences in the association of digital literacy with subjective health status and healthy lifestyle behaviors. All analyses were conducted using SAS software (version 9.4; SAS Institute), while statistical significance was set to $P<.05$. All analyses were weighted by NSOK to account for differences in survey completion rates across NSOK's survey regions.

Results

The participants' demographic and socioeconomic characteristics are summarized in Table 1. Of the 8664 study participants, all weighted by region, 55.9% (4649) had above-average scores in ITD usage, while 72.3% (6264) found it difficult to use information available online. The level of inconvenience of ITD usage was 4.2 out of 8.

Participants were mostly female (n=5077, 55.7%), aged 65‐69 years (n=3282, 35.8%), urban residents (n=6256, 76.1%), educated up to middle and high school (n=4571, 54.4%), married (n=4502, 59.7%), in household income level 4 (n=2169, 29.1%), and had 1‐2 chronic diseases (n=4936, 56.6%).

Table . Distribution of demographic and socioeconomic characteristics.

Variable	Total, n (%)[a]	Subjective health status			Healthy lifestyle behaviors		
		Healthy, n (%)[a]	Nonhealthy, n (%)[a]	*P* value[b]	Engagement, n (%)[a]	Nonengagement, n (%)[a]	*P* value[b]
Total	8664 (100)	4471 (51.1)	4193 (48.9)		3945 (46.9)	4719 (53.1)	
Degree of ITD[c] usage				<.001			<.001
Above average	4649 (55.9)	3038 (64.8)	1611 (35.2)		2329 (51.8)	2320 (48.2)	
Below average	4015 (44.1)	1433 (33.8)	2582 (66.2)		1616 (40.7)	2399 (59.3)	
Using information provided online				<.001			.05
Difficult	6264 (72.3)	2898 (45.6)	3366 (0.5)		2773 (46.1)	3491 (53.9)	
Nondifficult	2400 (27.7)	1573 (65.6)	827 (0.3)		1172 (48.9)	1228 (51.1)	
Inconvenience of ITD usage (score), mean (SD)	4.2 (2)	3.7 (1.9)	4.8 (2)	<.001	4.1 (2)	4.4 (2)	<.001
Gender				<.001			<.001
Male	3587 (44.3)	2097 (57)	1490 (43)		1427 (41.4)	2160 (58.6)	
Female	5077 (55.7)	2374 (46.5)	2703 (53.5)		2518 (51.3)	2559 (48.7)	
Age (years)				<.001			<.001
65 - 69	3282 (35.8)	2235 (69.3)	1047 (30.7)		1567 (48.2)	1715 (51.8)	
70 - 74	2193 (23.9)	1147 (53.9)	1046 (46.1)		1049 (48.6)	1144 (51.4)	
75 - 79	1668 (22.4)	640 (38.3)	1028 (61.7)		769 (48.8)	899 (51.2)	
80 - 85	1085 (12.9)	322 (27.2)	763 (72.8)		414 (40.9)	671 (59.1)	
≥85	436 (5)	127 (27)	309 (73)		146 (37.2)	290 (62.8)	
Residence				<.001			<.001
Urban (dong)	6256 (76.1)	3361 (52.8)	2895 (47.2)		3055 (50.4)	3201 (49.6)	
Rural (eup/myeon)	2408 (23.9)	1110 (45.9)	1298 (54.1)		890 (35.8)	1518 (64.2)	
Educational status				<.001			<.001
Elementary school or less	3609 (39.1)	1320 (34.3)	2289 (63.4)		1522 (42.5)	2087 (57.5)	
Middle and high school	4571 (54.4)	2790 (60.6)	1781 (39)		2123 (48.2)	2448 (51.8)	
University and above	484 (6.5)	361 (73.3)	123 (25.4)		300 (63)	184 (37)	

Variable	Total, n (%)[a]	Subjective health status			Healthy lifestyle behaviors		
		Healthy, n (%)[a]	Nonhealthy, n (%)[a]	*P* value[b]	Engagement, n (%)[a]	Nonengagement, n (%)[a]	*P* value[b]
Household type				<.001			.01
Single	2693 (19.8)	1117 (39.9)	1576 (60.1)		1160 (43)	1533 (57)	
Married couple	4502 (59.7)	2625 (57)	1877 (43)		2110 (47.8)	2392 (52.2)	
Cohabitant (nonspouse)	1469 (20.5)	729 (44.7)	740 (55.3)		675 (48.3)	794 (51.7)	
Household income				<.001			<.001
Level 1 (lowest)	2158 (20.7)	843 (38.4)	1315 (61.6)		799 (37.3)	1359 (62.7)	
Level 2	2171 (24.2)	965 (42.3)	1206 (57.7)		1044 (49.2)	1127 (50.8)	
Level 3	2166 (26)	1286 (58.2)	880 (41.8)		1020 (46.6)	1146 (53.4)	
Level 4 (highest)	2169 (29.1)	1377 (61.2)	792 (38.8)		1082 (52.1)	1087 (47.9)	
Number of chronic diseases				<.001			<.001
None	1533 (16.9)	1262 (82.1)	271 17.9		730 (48.2)	803 (51.8)	
1 - 2	4936 (56.6)	2696 (54.1)	2240 (45.9)		2160 (44.6)	2776 (55.4)	
3 or more	2195 (26.5)	513 (25)	1682 (75)		1055 (51.1)	1140 (48.9)	

[a]Weighted percentage.

[b]Rao-Scott chi-squared probability for comparison between (1) healthy and nonhealthy groups, and (2) groups engaging in healthy lifestyle behaviors and those not engaging in healthy lifestyle behaviors.

[c]ITD: IT device.

Of participants with above-average ITD usage, 64.8% (n=3038) perceived themselves as healthy and 51.8% (n=2329) engaged in healthy lifestyle behaviors. Of participants who reported difficulty using online information, 45.6% (n=2898) perceived themselves as healthy and 46.1% (n=2773) engaged in healthy lifestyle behaviors. Participants who perceived themselves to be healthy exhibited an ITD inconvenience score of 3.7 and those who engaged in healthy lifestyle behaviors exhibited a score of 4.1.

The association of digital literacy with subjective health status and healthy lifestyle behaviors is shown in Table 2. Compared with participants with below-average ITD usage, the odds of participants with above-average ITD usage were 1.73-times more likely to perceive themselves as healthy (adjusted odds ratio [aOR] 1.73, 95% CI 1.50 - 1.99, *P*<.001) and 1.51-times more likely to engage in healthy lifestyle behaviors (aOR 1.51, 95% CI 1.33 - 1.72, *P*<.001). Participants who did not find it difficult to use information available online had 1.41-times (aOR 1.41, 95% CI 1.24 - 1.61, *P*<.001) higher odds of perceiving themselves as healthy, compared with those who found it difficult. Higher ITD difficulty scores were also associated with 0.84-times lower odds of respondents perceiving themselves as healthy (aOR 0.84, 95% CI 0.82 - 0.87, *P*<.001) and 0.94-times lower odds of engaging in healthy lifestyle behaviors (aOR 0.94, 95% CI 0.92 - 0.97, *P*<.001).

Table . Digital literacy and its association with subjective health status and healthy lifestyle behaviors.

Variable	Perceived as healthy		Engaging in healthy lifestyle behaviors	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.73 (1.50-1.99)	<.001	1.51 (1.33-1.72)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.41 (1.24-1.61)	<.001	1.03 (0.92-1.16)	.59
Inconvenience of ITD usage (score)	0.84 (0.82-0.87)	<.001	0.94 (0.92-0.97)	<.001

[a]Adjusted for covariates including gender, age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref.: reference.

[e]Not applicable.

The association of digital literacy with subjective health status and healthy lifestyle behaviors by gender is shown in Tables 3 and 4. Both men and women with a higher than average degree of ITD usage were more likely to perceive themselves as healthy than those with an average or lower degree of ITD usage, and women in particular were nearly twice as likely (male aOR 1.64, 95% CI 1.34 - 1.99, *P*<.001; female aOR 1.91, 95% CI 1.60 - 2.27, *P*<.001). Similarly, people who did not experience any difficulties in using information provided online were more likely to perceive themselves as healthy than those who did experience difficulties, but this was more likely to be the case for men (male aOR 1.60, 95% CI 1.32 - 1.94, *P*<.001; female aOR 1.27, 95% CI 1.06 - 1.51, *P*=.01). Women with a higher degree of ITD usage were 1.54 times more likely to engage in healthy lifestyle behaviors than women with a lower degree of ITD usage, and this was slightly higher than for men (male aOR 1.47, 95% CI 1.20 - 1.81 *P*<.001; female aOR 1.54, 95% CI 1.30 - 1.82, *P*<.001). The less comfortable they were with using ITD, the less likely they were to perceive themselves as healthy (male aOR 0.84, 95% CI 0.80 - 0.88, *P*<.001; female aOR 0.85, 95% CI 0.81 - 0.88, *P*<.001) and to practice healthy lifestyles (male aOR 0.95, 95% CI 0.90 - 0.99, *P*=.02; female aOR 0.94, 95% CI 0.90 - 0.97, *P*<.001), with similar results for men and women.

Table . Digital literacy and its association with subjective health status by gender.

Variable	Perceived as healthy			
	Male		Female	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.64 (1.34-1.99)	<.001	1.91 (1.60-2.27)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.60 (1.32-1.94)	<.001	1.27 (1.06-1.51)	.01
Inconvenience of ITD usage (score)	0.84 (0.80-0.88)	<.001	0.85 (0.81-0.88)	<.001

[a]Adjusted for covariates including age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref.: reference.

[e]Not applicable.

Table . Digital literacy and its association with healthy lifestyle behaviors by gender.

Variable	Engaging in healthy lifestyle behaviors			
	Male		Female	
	aOR[a,b] (95% CI)	*P* value	aOR[a] (95% CI)	*P* value
Degree of ITD[c] usage				
Below average	Ref.[d]	—[e]	Ref.	—
Above average	1.47 (1.20-1.81)	<.001	1.54 (1.30-1.82)	<.001
Using information provided online				
Difficult	Ref.	—	Ref.	—
Nondifficult	1.08 (0.91-1.29)	.38	1.00 (0.85-1.18)	.96
Inconvenience of ITD usage (score)	0.95 (0.90-0.99)	.02	0.94 (0.90-0.97)	<.001

[a]Adjusted for covariates including age, residence, educational status, household type, household income, and number of chronic diseases.

[b]aOR: adjusted odds ratio.

[c]ITD: IT device.

[d]Ref: reference.

[e]Not applicable.

Discussion

Principal Findings and Comparison With Previous Work

This study examined the association between older people's digital literacy and their subjective health status and healthy lifestyle behaviors. Research has shown that older people with a high level of digital literacy, those who actively use information devices, and those who have easy access to online information are more likely to feel healthy and maintain healthy lifestyle behaviors. Conversely, those who experience difficulty accessing information online and are uncomfortable using information devices are less likely to engage in healthy lifestyle behaviors.

This study's findings are consistent with those of previous studies showing that both the use and amount of information affect subjective health status. In addition, whether older people own digital devices, how often they use them, and the extent of their internet activity are likely to affect their digital literacy and health [34-36]. According to a study in the United States, searching for and mastering relevant medical knowledge on the internet is beneficial for older adults with high blood pressure and heart disease, improving their health care and reducing disease rates [37]. In addition, using the internet could improve the health of older individuals by increasing their ability to exercise; for example, internet intervention programs helped them select exercise activities in the areas of endurance, flexibility, strengthening, and balance enhancement, which helped older adults increase their exercise frequency, thereby improving their physical health [38,39]. In China, older people with a high frequency of digital device usage experience a positive impact on their lifestyle, including improvements in diet, sleep, and exercise [15,40].

Empowering older people with a wide range of digital skills, from simple computer use to usage of social networks, emails, and online meetings, could enrich their lives, enhance their cognitive capabilities, and make them more productive and efficient members of society [41]. This could significantly contribute to the improvement of their perceived health status [42]. People who understand health information available online can use it effectively in their daily life to acquire valuable insights regarding their illnesses, increase their involvement as patient in the treatment process, and positively assess their subjective health status [43]. The internet provides a wealth of information, creating and increasing thematic content, especially for groups intimately associated with older adults, which helps maintain intergenerational relationships and promotes connected social capital [44]. With this social capital, older adults with a high level of digital literacy could fulfill their desire for belonging and love, which in turn could enable them to evaluate their health more positively [45].

However, unlike previous studies, this study found no difference in healthy lifestyle behaviors between those with and without difficulty using online information. Older people with high digital literacy have easier access to health information and use it effectively to positively impact their health care. Furthermore, they regularly access the internet for health information [27,35]. A previous study reported that adults aged 50 - 60 years who are not in good health access the internet for health care purposes because they want to be actively involved in health care decisions [46]. A few older individuals use emails and social networking sites to promote cancer screening to colleagues [47]. Online mediation, such as effectively managing hemoglobin A_{1c} levels in patients with diabetes through social networking site–based interventions, improved knowledge about the disease and had a positive impact on health behaviors [48,49].

Also in South Korea, Min [50] reported that the use of internet information decreased with age, regardless of physical and environmental accessibility, but this measure of use was comparable with this study's degree of ITD usage. Previous

studies have also highlighted the positive effects of online health information seeking psychological and physical domains, such as increased self-efficacy and improved health behaviors [51,52]. However, these studies were conducted before the COVID-19 pandemic, leaving a gap in understanding health behaviors during and after the pandemic.

The 2020 NSOK survey was conducted during the COVID-19 pandemic—a period marked by reduced outings among older adults due to social distancing policies and concerns about weakened immune systems. This context likely contributed to an overall decline in healthy lifestyle behaviors, particularly exercise, regardless of difficulties in using online information [23-25]. The reliance of Korean older adults on interpersonal sources for health information is deeply rooted in cultural values such as Confucianism and collectivism [53,54]. For health-related information, older adults in South Korea often rely on various sources, including health care professionals, mass media, local health centers, and family members. Particularly, adult children act as intermediaries in accessing and interpreting digital health information. In addition, a strong cultural trust in health care professionals reinforces a preference for direct consultations rather than independent online searches [55-57]. These sources play an important role in supporting health behaviors, even without the use of the internet, and help create an environment where information remains accessible despite low levels of online information use. In addition, older adults may have relied on household members or health center staff to mediate and explain online information, rather than accessing it directly [58-60]. This indirect approach could have resulted in an underestimation of the actual challenges older adults face when navigating online information.

Limitations

The present study had a few limitations. First, this is a part-sectional study. As we measured the variables at the time of the survey, we could only derive meaning in relation to the association between older adults' digital literacy and the causal relationship between their subjective health status and health practices. Second, as life expectancy increases, it is possible to divide the older adult population aged 85 years or older into different age groups. However, the NSOK is limited to detailed age-specific analysis. Nevertheless, this study used the latest data from national units, which ensured the representation of South Korea's older adults. In this study, digital literacy was divided into specific categories such as the degree of ITD usage, the difficulty of accessing online information, and the inconvenience of ITDs. This highlights the importance of linking older people's digital literacy to their subjective health status and health practices.

Conclusions

This study expands the existing knowledge on the subject by using a national-scale survey of South Korea's older adult population. To improve the subjective health and health practices of older adults, digital literacy education and related policies must be promoted. Furthermore, it is important to provide information education, so that older individuals can identify the required information, even among commercial and vague content. Health care and public institutions should provide information, booking systems, guidance, and other services to help older adults make use of these resources. We hope that this study will serve to improve older adults' quality of life and contribute to their health and welfare.

Data Availability

The data can be accessed from the Ministry of Health and Welfare and the Korea Institute for Health and Social Affairs (KIHASA) homepage. However, restrictions apply to data availability. The data were used with permission for the current study and, therefore, are not publicly available. The data are available from the corresponding author upon reasonable request and with prior permission from the KIHASA. Applications to use the National Survey of the Older Koreans (NSOK) data will be reviewed by the inquiry committee of research support, and once approved, raw data will be provided to the applicant for a fee.

Authors' Contributions

SYL was involved in study conception, statistical design, and manuscript drafting. YK and BK were involved in study conception, statistical design, data analysis, and manuscript drafting. SGL and SYJ oversaw the manuscript's development and helped revise it. THK was involved in study conception, statistical design, manuscript development, and manuscript revision. All the authors critically revised the manuscript for important intellectual content and approved its final version.

Conflicts of Interest

None declared.

References

1. Lee H, Lim JA, Nam HK. Effect of a digital literacy program on older adults' digital social behavior: a quasi-experimental study. Int J Environ Res Public Health 2022 Sep 29;19(19):12404. [doi: 10.3390/ijerph191912404] [Medline: 36231707]
2. 2023 population projections. Daejeon: Statistics Korea. 2023. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BPA003&conn_path=I2 [accessed 2024-06-13]
3. Murman DL. The impact of age on cognition. Semin Hear 2015 Aug;36(3):111-121. [doi: 10.1055/s-0035-1555115] [Medline: 27516712]

4. Checton MG, Greene K. Elderly patients' heart-related conditions: disclosing health information differs by target. Psychol Health Med 2015;20(5):594-604. [doi: 10.1080/13548506.2014.986141] [Medline: 25434731]
5. Chaudhuri S, Le T, White C, Thompson H, Demiris G. Examining health information-seeking behaviors of older adults. Comput Inform Nurs 2013 Nov;31(11):547-553. [doi: 10.1097/01.NCN.0000432131.92020.42] [Medline: 23974574]
6. An S, Lim Y, Chung S. Psychological wellbeing through health information behavior: effects of the elderly's digitigal literacy. Korea Gerontol Soc 2020 Oct 1;40(5):833-854. [doi: 10.31888/JKGS.2020.40.5.833]
7. Global strategy on digital health 2020–2025. : World Health Organization; 2021 URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-05-30]
8. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
9. Schäffer B. The digital literacy of seniors. Research in Comparative and International Education 2007 Mar;2(1):29-42. [doi: 10.2304/rcie.2007.2.1.29]
10. Gilster P. Digital Literacy: Wiley Computer Publishing; 1997.
11. Leahy D, Dolan D. Digital literacy: a vital competence for 2010. Presented at: IFIP International Conference on Key Competencies in the Knowledge Society; Sep 20-23, 2010; Brisbane, Australia. [doi: 10.1007/978-3-642-15378-5_21]
12. Park H, Jin M, Park J, Lim K. The analysis of national and international curriculum for digital literacy education. J Educ Cult 2021;27(5):75-101. [doi: 10.24159/joec.2021.27.5.75]
13. Cockerham WC, Sharp K, Wilcox JA. Aging and perceived health status. J Gerontol 1983 May;38(3):349-355. [doi: 10.1093/geronj/38.3.349] [Medline: 6841932]
14. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord 2019 Feb 15;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
15. Zhou Y, Bai Y, Wang J. The impact of internet use on health among older adults in China: a nationally representative study. BMC Public Health 2024;24(1):1065. [doi: 10.1186/s12889-024-18269-4]
16. Lee JY. A study on the use of information and social computing service by the elderly. Journal of the Korean Society for information Management 2012 Mar 30;29(1):375-393. [doi: 10.3743/KOSIM.2012.29.1.375]
17. Miller LMS, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health 2012 Apr;24(3):525-541. [doi: 10.1177/0898264311428167] [Medline: 22187092]
18. The report on the digital divide [Article in Korean]. : Ministry of Science and ICT, National Information Society Agency; 2022 URL: https://www.nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=81623&bcIdx=25353&parentSeq=25353 [accessed 2025-06-04]
19. Ahn JI, Seo YK. An analysis of sub-factors of digital media literacy divide: focusing on the effects of generation and economic status. Journal of Digital Convergence 2014 Feb 28;12(2):69-78. [doi: 10.14400/JDC.2014.12.2.69]
20. Choi HI, Song I. The mediating effect of self-efficacy between the elderly's digital information literacy and life satisfaction. J Korea Acad Ind Coop Soc 2020;21(6):246-255. [doi: 10.5762/KAIS.2020.21.6.246]
21. Kim M, Kim Y, Um S. A study about the types of digital information utilization and life satisfaction of middle-aged and the older people. J Korean Assoc Reg Inf Soc 2020;21:51-74. [doi: 10.22896/karis.2020.23.1.003]
22. Park S, Kim B. Predictors of internet use among older adults with diabetes in South Korea: survey study. JMIR Med Inform 2020 Dec 23;8(12):e19061. [doi: 10.2196/19061] [Medline: 33277232]
23. Machado CLF, Pinto RS, Brusco CM, Cadore EL, Radaelli R. COVID-19 pandemic is an urgent time for older people to practice resistance exercise at home. Exp Gerontol 2020 Nov;141:111101. [doi: 10.1016/j.exger.2020.111101] [Medline: 33049354]
24. Roschel H, Artioli GG, Gualano B. Risk of increased physical inactivity during COVID-19 outbreak in older people: a call for actions. J Am Geriatr Soc 2020 Jun;68(6):1126-1128. [doi: 10.1111/jgs.16550] [Medline: 32392620]
25. Kim J, Kim Y, Ha J. Changes in daily life during the COVID-19 pandemic among South Korean older adults with chronic diseases: a qualitative study. Int J Environ Res Public Health 2021 Jun 24;18(13):6781. [doi: 10.3390/ijerph18136781] [Medline: 34202534]
26. Kim S, Choi H. Still hungry for information: information seeking behavior of senior citizens in South Korea. Presented at: iConference 2014 Proceedings; Mar 4-7, 2014; Berlin, Germany p. 889-894. [doi: 10.9776/14300]
27. Levy H, Janke AT, Langa KM. Health literacy and the digital divide among older Americans. J Gen Intern Med 2015 Mar;30(3):284-289. [doi: 10.1007/s11606-014-3069-5] [Medline: 25387437]
28. Mubarak F, Suomi R. Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide. INQUIRY 2022;59:469580221096272. [doi: 10.1177/00469580221096272] [Medline: 35471138]
29. Lee YK, Kim S, Hwang N, Lim J, Joo B, Namgung E, et al. 2020 National Survey of Older Koreans. : Korea Institute for Health and Social Affairs; 2020.
30. Choi Y. Is subjective health reliable as a proxy variable for true health? A comparison of self-rated health and self-assessed change in health among middle-aged and older South Koreans. Health and Social Welfare Review 2016 Dec;36(4):431-459. [doi: 10.15709/hswr.2016.36.4.431]

31. Benyamini Y, Idler EL, Leventhal H, Leventhal EA. Positive affect and function as influences on self-assessments of health: expanding our view beyond illness and disability. J Gerontol B Psychol Sci Soc Sci 2000 Mar;55(2):107-116. [doi: 10.1093/geronb/55.2.p107] [Medline: 10794189]
32. Heller DA, Ahern FM, Pringle KE, Brown TV. Among older adults, the responsiveness of self-rated health to changes in Charlson comorbidity was moderated by age and baseline comorbidity. J Clin Epidemiol 2009 Feb;62(2):177-187. [doi: 10.1016/j.jclinepi.2008.05.009] [Medline: 18722090]
33. Peel NM, McClure RJ, Bartlett HP. Behavioral determinants of healthy aging. Am J Prev Med 2005 Apr;28(3):298-304. [doi: 10.1016/j.amepre.2004.12.002] [Medline: 15766620]
34. Cherid C, Baghdadli A, Wall M, et al. Current level of technology use, health and eHealth literacy in older Canadians with a recent fracture-a survey in orthopedic clinics. Osteoporos Int 2020 Jul;31(7):1333-1340. [doi: 10.1007/s00198-020-05359-3] [Medline: 32112118]
35. Tennant B, Stellefson M, Dodd V, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70. [doi: 10.2196/jmir.3992] [Medline: 25783036]
36. Papp-Zipernovszky O, Horváth MD, Schulz PJ, Csabai M. Generation gaps in digital health literacy and their impact on health information seeking behavior and health empowerment in Hungary. Front Public Health 2021;9:635943. [doi: 10.3389/fpubh.2021.635943] [Medline: 34055714]
37. Meischke H, Eisenberg M, Rowe S, Cagle A. Do older adults use the Internet for information on heart attacks? Results from a survey of seniors in King County, Washington. Heart Lung 2005;34(1):3-12. [doi: 10.1016/j.hrtlng.2004.06.006] [Medline: 15647729]
38. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based Intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013;15(2):e19. [doi: 10.2196/jmir.2158]
39. Guo E, Li J, Luo L, Gao Y, Wang Z. The effect and mechanism of internet use on the physical health of the older people-empirical analysis based on CFPS. Front Public Health 2022;10:952858. [doi: 10.3389/fpubh.2022.952858] [Medline: 36339140]
40. Cui K, Zou W, Ji X, Zhang X. Does digital technology make people healthier: the impact of digital use on the lifestyle of Chinese older adults. BMC Geriatr 2024 Jan 22;24(1):85. [doi: 10.1186/s12877-023-04651-1] [Medline: 38254001]
41. Martínez-Alcalá CI, Rosales-Lagarde A, Pérez-Pérez YM, Lopez-Noguerola JS, Bautista-Díaz ML, Agis-Juarez RA. The effects of Covid-19 on the digital literacy of the elderly: norms for digital inclusion. Front Educ 2021;6. [doi: 10.3389/feduc.2021.716025]
42. Prieto-Flores ME, Moreno-Jiménez A, Fernandez-Mayoralas G, Rojo-Perez F, Forjaz MJ. The relative contribution of health status and quality of life domains in subjective health in old age. Soc Indic Res 2012 Mar;106(1):27-39. [doi: 10.1007/s11205-011-9791-z]
43. Song JH, Shin SJ. The effects of e-Health literacy and subjective health status on health-seeking behaviors of elderly using the internet in the community. J Digit Converg 2020;18(1):321-332. [doi: 10.14400/JDC.2020.18.1.321]
44. Lu H, Kandilov IT. Does mobile internet use affect the subjective well-being of older Chinese adults? An instrumental variable quantile analysis. J Happiness Stud 2021 Oct;22(7):3137-3156. [doi: 10.1007/s10902-021-00365-6]
45. Yang Y, Zeng D, Yang F. Internet use and subjective well-being of the elderly: an analysis of the mediating effect based on social capital. Int J Environ Res Public Health 2022 Sep 24;19(19):12087. [doi: 10.3390/ijerph191912087] [Medline: 36231387]
46. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 2;16(12):e262. [doi: 10.2196/jmir.3706] [Medline: 25470306]
47. Cutrona SL, Roblin DW, Wagner JL, et al. Adult willingness to use email and social media for peer-to-peer cancer screening communication: quantitative interview study. JMIR Res Protoc 2013 Nov 28;2(2):e52. [doi: 10.2196/resprot.2886] [Medline: 24287495]
48. Yaagoob E, Hunter S, Chan S. The effectiveness of social media intervention in people with diabetes: an integrative review. J Clin Nurs 2023 Jun;32(11-12):2419-2432. [doi: 10.1111/jocn.16354] [Medline: 35545822]
49. Kim JN, Lee S. Communication and cybercoping: coping with chronic illness through communicative action in online support networks. J Health Commun 2014;19(7):775-794. [doi: 10.1080/10810730.2013.864724] [Medline: 24559492]
50. Min Y. The digital divide among internet users: an analysis of digital access, literacy, and participation. J Commun Res 2011 Feb;48(1):150-187. [doi: 10.22174/jcr.2011.48.1.150]
51. Kim S, Love F, Quistberg DA, Shea JA. Association of health literacy with self-management behavior in patients with diabetes. Diabetes Care 2004 Dec;27(12):2980-2982. [doi: 10.2337/diacare.27.12.2980] [Medline: 15562219]
52. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. Health Educ Behav 2015 Apr;42(2):202-209. [doi: 10.1177/1090198114547815] [Medline: 25156311]
53. Koh EK, Koh CK. Caring for older adults: the parables in Confucian texts. Nurs Sci Q 2008 Oct;21(4):365-368. [doi: 10.1177/0894318408324320] [Medline: 18953016]

54. Palley HA, Usui C. Social policies for the elderly in the Republic of Korea and Japan: a comparative perspective*. Soc Policy Adm 1995 Sep;29(3):241-257. [doi: 10.1111/j.1467-9515.1995.tb00467.x]
55. Koo S, Cho J, Ahn E, et al. Use of health information among older adults: trust and source of health information by gender and region. Health and Social Welfare Review 2016 Jun;36(2):454-479. [doi: 10.15709/hswr.2016.36.2.454]
56. Jung WS, Kang HG, Han S, Kim E. Recognition and use of health information for preliminary elderly and elderly people. Journal of Digital Convergence 2021;19(6):419-427. [doi: 10.14400/JDC.2021.19.6.419]
57. Chang SJ, Lee KJ, Kim IS, Lee WH. Older Korean people's desire to participate in health care decision making. Nurs Ethics 2008 Jan;15(1):73-86. [doi: 10.1177/0969733007083936] [Medline: 18096583]
58. Kim DJ, Lee YS, Jeon ER, Kim KJ. Present and future of AI-IoT-based healthcare services for senior citizens in local communities: a review of a South Korean Government digital healthcare initiatives. Healthcare (Basel) 2024;12(2):281. [doi: 10.3390/healthcare12020281]
59. Jang Y, Jung J, Park NS, Kim MT, Chung S. COVID-19 health information sources and their associations with preventive behaviors: a typological study with older residents in Seoul, South Korea. PLoS One 2023;18(8):e0290824. [doi: 10.1371/journal.pone.0290824] [Medline: 37651388]
60. Weber W, Reinhardt A, Rossmann C. Lifestyle segmentation to explain the online health information-seeking behavior of older adults: representative telephone survey. J Med Internet Res 2020 Jun 12;22(6):e15099. [doi: 10.2196/15099] [Medline: 32530433]

Abbreviations

aOR: adjusted odds ratio
IRB: Institutional Review Board
ITD: IT device
KIHASA: Korea Institute for Health and Social Affairs
NSOK: National Survey of the Older Koreans

Edited by J Wang; submitted 01.08.24; peer-reviewed by J Kim, M Tomita; revised version received 02.03.25; accepted 11.04.25; published 05.06.25.

Please cite as:
Lee SY, Kim Y, Kim B, Lee SG, Jang SY, Kim TH
Digital Literacy and Its Association With Subjective Health Status and Healthy Lifestyle Behaviors Among Korean Older Adults: Cross-Sectional Study
JMIR Aging 2025;8:e64974
URL: https://aging.jmir.org/2025/1/e64974
doi:10.2196/64974

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org